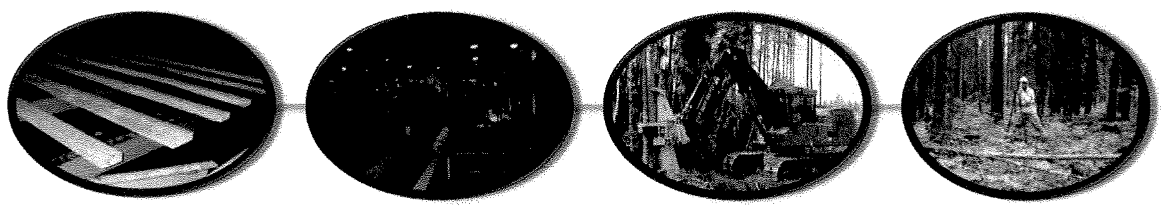
2012-11-10

Pre-harvest lumber value recovery modeling: integrating sawline laser-scanning with an enhanced forest inventory

Shorthouse, Kevin D.

PRE-HARVEST LUMBER VALUE RECOVERY MODELING: INTEGRATING SAWLINE LASER-SCANNING WITH AN ENHANCED FOREST INVENTORY



by

Kevin D. Shorthouse

A Master's Thesis Submitted in
Partial Fulfillment of the Requirements for the
Degree of Master of Science in Forestry

Faculty of Natural Resources Management
Lakehead University
Thunder Bay, Ontario
February 7th, 2012

LIBRARY RIGHTS STATEMENT

A CAUTION TO THE READER

This M.Sc.F. thesis has been through a formal process of review and comment by at least three faculty members and an external examiner. It is made available for loan by the Faculty of Natural Resources Management for the purpose of advancing the practice of professional and scientific forestry.

The reader should be aware that opinions and conclusions expressed in this document are those of the student and do not necessarily reflect the opinions of either the thesis supervisors, the faculty or Lakehead University.

ABSTRACT

Shorthouse, K.D. 2012. Pre-harvest lumber value recovery modeling: integrating sawline laser-scanning with an enhanced forest inventory.

Keywords: lumber value recovery, boreal, stud sawmill, log scanning, enhanced forest inventory, supply chain, principal component analysis, multiple regression.

Supply chain management research in forestry is becoming increasingly important in the forest sector worldwide. Predicting lumber value recovery from pre-harvest forest inventory forms an important aspect of forestry supply chain management research. The purpose of this research was to conduct a case study for tracking a stud sawmill's supply chain in Ontario's boreal forest from pre-harvest forest inventory through to green lumber value recovery (GLVR). The specific objectives of the research were to: (i) compile a detailed dataset along a stud sawmill value chain; (ii) explore the similarity of variables existing in log- and tree-level datasets; (iii) extract Principal Components that describe the variance and collinearity of the variables found in the log- and tree-level datasets; (iv) build regression models for a stud sawmill for predicting GLVR; and (v) compare and assess model performance in predicting GLVR.

The log-level analysis was conducted by measuring 16,565 log profiles and their GLVR at the sawmill. Of the 7 external log variables measured, 6 variables yielded two principal components: log size and log form. For the tree-level analysis, 101 standing trees were measured in the pre-harvest timber cruise. Of the 29 external tree variables measured, 10 variables yielded three components: tree size, tree form, and tree branchiness. The multiple regression analysis using the principal components found that non-linear exponential function produced the strongest log-level and tree-level models for predicting the GLVR. The study demonstrated that an enhanced pre-harvest forest inventory can be easily integrated with sawline laser-scanning sawmill data to estimate the GLVR prior to harvesting.

CONTENTS

TABLES

FIGURES

ACKNOWLEDGEMENTS

I would firstly like to thank my research committee for their guidance and support throughout this research project. My co-supervisor, Prof. Reino Pulkki, helped create the project and provided invaluable instruction along the way. My other co-supervisor, Prof. Chander Shahi, provided expert modeling advice and guided me through the challenging analysis. Committee member, Prof. Mathew Leitch, facilitated the partnership with Resolute Forest Products sawmill and provided a valuable addition to the supply chain dataset by conducting internal wood properties testing within his research group. I would also like to extend my thanks to Prof. Jean-Marc Frayret for agreeing to be the external examiner. His comments and expert knowledge in forest sector supply chain research were invaluable.

I would like to thank those who helped me in the field to collect data for the project. Those who helped me included graduate students: Serge Laforest, Brent Forbes, Steve Hosegood and Bedarul Alam. I also had assistance from two of my professors: Prof. Reino Pulkki and Prof. Mathew Leitch.

I am indebted to our industry partners Resolute Forest Products, Marcri Logging and CCS Central Computer Services; without their support this project would not have been possible. Specifically I would like to thank Roger Leclerc from the Resolute sawmill who was behind the project from its inception and provided exceptional leadership and expert advice in carrying out the project. I would also like to specifically thank Bill Smith from Resolute's woodlands department, who facilitated study site selection and coordination with Marcri Logging. NSERC and CCS Central Computer Services also helped fund the research project and CCS provided technological support for which I am grateful.

Finally I would like to thank my family for their continual support and care throughout this challenging research project. I have had the great privilege of marrying my wife, Brandi Shorthouse, while completing my Master's and she has been a continual source of love and encouragement. Brandi also played a valuable editing role. When I was in distress she always helped me see that the end was in sight and encouraged me to persevere until I finished.

1. INTRODUCTION

1.1 BACKGROUND

Supply chain management (SCM) research in forestry is becoming increasingly important in the forest sector worldwide (Sjostrom and Rask 2001). Synonyms of this research term include: total production management, value chain management, logistical management, holistic approach and total system costing (Ellram 1991; Cooper *et al.* 1997; Christopher 1998; Smith 1999; Mentzer *et al.* 2001; Pulkki 2004; Haartveit *et al.* 2004; D'Amours *et al.* 2011). Although individual definitions differ substantially, most agree with Christopher (1998) that "the ultimate purpose of any logistics system is to satisfy customers" and that customer satisfaction is one of the central issues of SCM. Furthermore, SCM is the proper planning, organization, development, coordination, steering and control of inter- and intra-organizational processes in a holistic manner including exchanges of information, materials, funds, and product development and marketing activities along the entire production chain of a business process (Mattson 2000; Haartveit *et al.* 2004).

SCM terminology first emerged approximately 25 years ago (Ganeshan and Harrison 1995), but it is only since 1990 that SCM has been applied broadly in the literature to describe a complete business system, mapping information and inventory from raw materials to delivery of end products to customers (Pulkki 2004). In recent years, the term has evolved further to what is known as value chain management (VCM), value chain optimization (VCO) or simply 'value chains'. Within research on value chains there is an increased emphasis on the customer and a holistic approach to the existing SCM. Value means the optimization of a combinatorial set of values including profit, cost, service levels, productivity, employment and environment values. Value also includes production and distribution capacities, all affecting the capacity to create value. The 'chain' in value chain can be described as a series of links that integrate to make an enterprise. For example, the broad chain links in a forest enterprise will include: the forest resource, procurement logistics, production and delivery to customers. Additionally, forest value chains can be modeled over multiple time horizons extending from short-term local harvesting decisions to long-term regional forest strategies (D'Amours *et al.* 2010).

The supply chains that demonstrate the ability to track and optimize forest fibre resources from the forest to final product and eventually to the customer will determine the groups who will remain competitive in the future (Poirier 1999). It is crucial for companies and industrial clusters in the forest sector to apply principals of supply chain management in order to maximize profit and minimize inventory, as well as information integration along each step of the supply chain (Mattson 1999; Pulkki 2004; Haartveit *et al.* 2004; Alam and Pulkki 2009).

The forest sector is a complex supply chain with a high level of uncertainty in fibre forecasts from forest to manufacturer. Historical, and more recent studies, have demonstrated that accurate estimates of fibre volume, quality and potential products from the forest inventory, are the foundation for optimizing forest product value chains (Deadman *et al.* 1979; Pulkki 1990; Deadman 1990; Pulkki 1991; Middleton *et al.* 1995;

Uusitalo 1997; Middleton *et al.* 2003; Maltamo *et al.* 2003; Maltamo *et al.* 2006; Liu *et al.* 2007a; Liu *et al.* 2007b; Li 2009; Middleton and Zhang 2009; Groot and Pitt 2010).

In VCM, the focus is increasing towards customer satisfaction through accurate and efficient order – or available-to-promise (ATP) – fulfillment. Available-to-promise fulfillment is a business function that provides the most efficient response to customer orders (Ball *et al.* 2004; Fleishmann and Meyr 2004). To maximize profits, ATP fulfillment is becoming increasingly important in an increasingly market-led economic environment (D'Amours *et al.* 2008). In reality, production economies are always market-led and consumer-driven; however, in light of recent global economic challenges, particularly for the forest industry, conventional value chain models and long-time business practices are being challenged.

Traditionally, the Canadian forest industry has planned production based on broad estimates of wood volume by species and relied on log yards with large inventories to allocate the appropriate sized logs and wood quality to fulfill production orders (Dramm 2004; Gaudreault *et al.* 2009). Through reliance on a volume-oriented business model with large inventories geared toward the production of commodity products, the Canadian forest industry is no longer sustainable (D'Amours *et al.* 2010). The strategic and tactical analyses done by the forest industry have been aimed at total wood volumes with little regard to the requirements of the entire supply chain. In an increasingly competitive market place with increasing production costs, fibre quality and intrinsic value need to be classified as equal if not more important to fibre volume. For a sawmill, tree or log piece size combined with its form and fibre quality will predict product value recovery (Zhang and Gingras 1999). Therefore, a paradigm shift needs to occur in our forest allocation planning to include wood quality characteristics (e.g., log piece size and quality) in order to plan for wood value and not merely volume (Pulkki 2004). Additionally, forest product manufacturers in Canada have traditionally focused the majority of production on commodity products and have primarily sold to one customer: the United States. Due to a number of economic and global competition factors the Canadian forest sector is restructuring, which includes a broader range of higher-value products and a larger customer base.

An important part in the restructuring process of the Canadian forest sector is to understand and model its supply chains from pre-harvest inventory to market-ready products (Frayret *et al.* 2007; Groot and Pitt 2010). Pre-harvest inventory is a specialized form of sampling and statistical analysis used to collect information about the species present in the forest, how abundant they are, how they are distributed in the area, and the range of sizes and size class distribution. This information is then used in the appraisal of forest growing stock for specific products before harvesting (Avery and Burkhart 1983; Davis and Johnson 1987; Nieuwenhuis 2002; Higman *et al.* 2005). In a pre-harvest forest inventory, there can be a wide range in the level of detail and variables measured; this depends on the level and amount of information managers require to make tactical and strategic decisions to best satisfy the management objectives for any given forest management unit (Higman *et al.* 2005; Li 2009). An enhanced pre-harvest inventory refers to a more detailed kind of pre-harvest inventory, which can include detailed specifications of the forest resource to predict product recovery; this includes stem profile, quality and internal fibre characteristics (Pulkki 1990; Deadman 1990; Pulkki 1991; Uusitalo 1997; Maltamo *et al.* 2003; Maltamo *et al.* 2006; Liu *et al.* 2007a, Liu *et al.* 2007b; Li 2009; Groot and Pitt 2010).

As forest supply chains become more competitive, the ability to utilize an enhanced pre-harvest forest inventory to predict different characteristics of product recovery will equip managers with better decision-making tools to optimize the available resources for customer satisfaction. Although widely researched globally, models of how an enhanced pre-harvest inventory can be used to predict product recovery has not been extensively studied in Canada. There is a pivotal knowledge and technical innovation gap between how individual mills can apply the broad body of research to create tools specific to their supply chain that can simulate product recovery from pre-harvest inventory.

Additionally, there is increasing pressure on forest enterprises to tighten the supply chain and reduce inventory costs. For example, mills running with a tighter supply chain will usually hold only 3 to 4 days of inventory, irrespective of harvesting plans that can change monthly or even weekly depending on the type of product required to fulfill customer orders (D'Amours *et al.* 2008; Gaudreault *et al.* 2009; D'Amours *et al.* 2011). Tighter supply chains can become problematic if not coupled with agile logistics-management systems that can efficiently adapt to changes in the market place (Frayret 2001; D'Amours *et al.* 2010). Benefits to a tighter supply chain with agile manufacturing logistics include reduced inventory-holding costs, reduced wood degradation and higher sales potential (Pulkki 1991; Frayret 2001; D'Amours *et al.* 2010).

However, in order to plan for a tighter supply chain, decision support tools (DST) need to be developed that balance efficient use of the harvesting and transportation operations with the manufacturing goals. These tools are usually GIS-based platforms that can use pre-harvest inventory to forecast wood size, species and quality (Sjostrom and Rask 2001; Uusitalo 2005). These GIS-based DSTs assist managers in deciding which forest compartments should be harvested to satisfy customer demand. Developing DSTs for the optimal allocation of forest sections helps to achieve the ultimate goal of value maximization over the entire value chain. Put succinctly, the goal of value chain optimization is to deliver the right species of wood of appropriate size and quality specification to the right mill at the right time to fulfill market demand (Sjostrom and Rask 2001; Uusitalo 2005; D'Amours 2008).

Product recovery simulation supports the strategic allocation of forest resources to satisfy specific customer orders in a forest value chain. Many sawmills – and other forest product manufacturers – worldwide are now using sophisticated lumber recovery simulation tools, (e.g., Halco Software Systems) and statistical techniques (e.g., Principal Component Analysis [PCA] and multiple regression) that utilize enhanced pre-harvest inventories to model sawmill product value recovery pre-harvest (Steele 1984; Liu *et al.* 1989; Shi *et al.* 1990, Wagner and Taylor 1993; Roos *et al.* 2000; Via *et al.* 2003; Middleton *et al.* 2003; Liu *et al.* 2007a; Liu *et al.* 2007b). These types of product recovery simulators can be used in creating geo-database decision support tools to conduct short and medium-term operational planning in the forest to optimize the wood mix being transported to one or more mills to fulfill specific orders. In Canada, modeling product recovery using pre-harvest data is still in the research phase or early implementation for progressive industry partners, and thus, is an important knowledge gap both in the research literature and for the Canadian forest sector.

1.2 OBJECTIVES

The purpose of the study is to test whether an enhanced forest inventory can be integrated with sawmill production data to model green lumber value recovery (GLVR) prior to forest harvesting.

The specific objectives for this study are: to (i) compile a detailed dataset along a stud sawmill value chain in Ontario's Boreal forest from pre-harvest to green lumber value production; (ii) determine the levels of similarity existing in log- and tree-level datasets; (iii) extract Principal Components that describe the majority of variance and collinearity of the variables found in the log- and tree-level datasets; (iv) build GLVR models for a stud sawmill using PCA regression scores; and (v) to compare and assess model performance in predicting GLVR.

1.3 SCOPE

The dataset described in the Methods contains a huge amount of data. For this reason the scope of the study was limited to the pre-harvest and saw-line recovery datasets. The dataset collected remains of high-value with potential for a number of future research projects under the topic of forest value chain research.

Additionally, the study has some limitations. Firstly, due to a small number of trial sites focusing on jack pine (*Pinus banksiana* Lamb.), the recovery models created could not be validated and should be applied with discretion for jack pine trees and logs with size and form characteristics outside the range of the trees measured. Additionally, the thesis models green lumber, not kiln-dried and planed lumber value recovery. GLVR was modeled due to constraints in the sawmill w did not allow tracking of individual lumber pieces from each log through to the planer mill and final grading. Also, the GLVR estimates from pre-harvest inventory are simulated estimates, as it was not feasible to track each tree measured through to lumber conversion. Finally, the Comact laser scanners and optimizers at the sawmill did not take into account log knot characteristics, which – we know from the literature review – have a significant impact on the lumber value recovery of jack pine.

2. LITERATURE REVIEW

2.1 THE PARADIGM SHIFT IN THE CANADIAN FOREST SECTOR

The forest industry is a significant contributor to Canada's economy (3% GDP), and annual exports of lumber, pulp and paper are estimated at $45 billion, contributing 60% to the trade surplus (FPAC 2007). The forest industry is also a major source of employment, contributing 240,000 direct jobs across the country (FPAC 2012). However, the industry has recently faced unprecedented global economic challenges, including international competition, a strong Canadian dollar, a weakened US economy, obsolete production technologies, increasing cost of energy, trade dispute with the US, and reduced access to capital due to poor past financial performance (FPAC 2007; D'Amours *et al.* 2010).

Canada's domestic environment has contributed to the non-competitiveness as a result of high-energy costs, high delivered wood costs, relatively high labour costs and outdated mills. Another contributing factor is the lack of research and development (R&D) leadership from government and industry (Mandel-Campbell 2007). Canada's traditional lumber, pulp and paper industries have failed to invest in R&D related to new or innovative forest products that could fulfill previously unmet markets by capitalizing on regional competitive strengths (FPAC 2007; Gingras 2011). Furthermore Canada has failed to develop a global forest product strategy (Emmett 2005; FPAC 2007; Mandel-Campbell 2007).

Traditionally, Canada's forest industry maintained a push supply model, where manufacturers focused on producing commodity products, such as pulp, paper and lumber. Under this model, the industry focused on inventory production and assumed growing markets would be available to buy the product (Buongiorno 2003; Beaudoin *et al.* 2007). In light of the growing non-competiveness of our commodity producers, there is demand for a new model. This model is customer oriented (pull-supply model); the process starts with a market demand for specific forest product attributes (Beaudoin *et al.* 2007). The pull model supports diversified and value-added forest products that are linked directly to the attributes found at the tree level (Beaudoin *et al.* 2007; Groot and Pitt 2010). In order to fully utilize the pull model, forest value chains are being mapped from consumer, through the manufacturing process, and back to forest management and operations where particular wood fiber and characteristics are desired (D'Amours 2010; Groot and Pitt 2010).

As a result of these challenges, there have been efforts to find a new policy direction in Canada's Forest Sector (Mackenzie and Bruemmer 2009). One of the foundational points in the policy shift is a paradigm shift away from an exclusively volume and commodity product based industry to a value oriented industry, where value is a function of fibre volume, form and quality (FPAC 2007; Lazar 2007). By relying on large inventories in the production of commodity products, the forest inventory paradigm has been dominated by a volume-oriented management style. The shift from a volume-oriented commodity model to a quality-oriented value-added model is not an easy shift and requires significant R&D in our forest value chains from tree and fibre characteristics through to intensive market research. In order to gain a better understanding of Canada's forest resource value, there is a need to understand which fibre attributes and tree characteristics are significant to forecast product recovery. Additionally there is need to develop GIS-based inventory systems that can show how value characteristics vary across species and geography (Groot and Pitt 2010). As paradigms in policy and breakthroughs in research bring innovation to the forest sector, Canada will create a climate of investment in the forest economy and secure its future prosperity (Emmett, 2005).

The policy move towards an enhanced forest resource inventory using the latest technology is supported by research both abroad and now by research executed in Canada. New Zealand's forest business model describes their forest estate as a warehouse with a detailed inventory. The concept of treating the commercial forest areas as warehouses requires detailed information of the types of wood/fibre and wood/fibre location, and how to best allocate these wood/fibre types for optimal value creation (Tomppo *et al.* 1999; Goulding 2000; Goulding *et al.* 2000). Creating optimal value requires market-driven harvest planning in order to drive management decisions at the strategic, tactical and operational levels (Beaudoin *et al.* 2007). More recent research in Canada is exploring how to integrate new technology (e.g. lidar, high-resolution imagery, etc.) to obtain an enhanced forest inventory to model forest product recovery (Groot and Pitt 2010).

2.2 RESEARCH AND DEVELOPMENT CLUSTERS

Several research clusters have been formed to find solutions to the challenges facing the Canadian forest sector. These research clusters combine researchers from government, industry and universities. FPInnovations is one of Canada's leading national research hubs for creating specific research clusters (or groups) and addresses specific needs of the forest industry. FPInnovations' mandate is to optimize the forest sector value chain by capitalizing on uniquely Canadian fibre attributes and to develop new products and market opportunities within a framework of environmental sustainability (Gingras 2011). FPInnovations has four research groups involved in forestry value chain research including the: Forest Operations Division, Wood Products Division, Pulp and Paper Division and Canadian Wood Fibre Centre (CWFC) (FPInnovations 2011).

Additionally, the Natural Sciences and Engineering Research Council of Canada (NSERC), FPInnovations and Natural Resources Canada (NRCan) have developed a networking hub called the NSERC Forest Sector R&D Initiative; developed to establish, coordinate and fund strategic networks to bring innovative solutions to the Canadian forest sector. There are eight strategic networks currently working under the NSERC Forest Sector R&D Initiative including: Green Fibre Network, NEWBuildS, VCO Network, Bioconversion, Sentinel, Papier, ArboraNano and ForValueNet (NSERC 2011, FPInnovations 2011). An important characteristic of these networks is the cooperative focus between groups (i.e., universities, government and industry). These networks develop cohesive research strategies that coordinate individual research topics, supply funding, provide an integrative platform for independent groups and have technology transfer to the industry as a major goal.

Strategic research clusters under FPInnovations and NSERC are taking a market-focused strategy and using new technologies to find innovative solutions to induce global competiveness to the Canadian forest sector (D'Amours *et al.* 2010; Gingras 2011). There has been a renewed interest in market research, as well as the implementation of marketing principles for the forest industry, towards "being customer focused, but not customer led" (Slater and Narver 1998; FPAC 2007; D'Amours *et al.* 2010; Gingras 2011). Forest product industries are now approached as Value Creation Networks (VCNs), which includes all the companies and business units involved in the procurement, production of a given product and its distribution to the market (Velde 2006; Beaudoin *et al.* 2007; Frayret *et al.* 2007; D'Amours *et al.* 2008). These VCNs are being transformed to create new products and to target new markets. An important component to include within this definition is the information flow, transfer and integration between components in the VCN. The goal of the VCN is to optimize value for the entire group, not simply individual components (Cambiaghi *et al.* 2008; D'Amours *et al.* 2008).

The Canadian forest sector requires innovation to become globally competitive and overcome the current industry crisis. The prevailing government strategy set to deal with the current fallout is funding R&D to re-examine the entire supply chain from tree to customer. FORAC and CIRRELT are two research groups that have emerged over the past decade as leaders in discovering innovative solutions for the Canadian forest sector. The FORAC research cluster is a centre of expertise for advancement of the forest products industry (FORAC 2011). FORAC's focus is to develop global quality research in the fields of integration and optimizations of the value creation network in the forest products industry (FORAC 2011). CIRRELT is a Quebec-based research cluster focused on optimizing the usage of networks (CIRRELT 2011). FORAC and CIRRELT researchers have been leading-partners in the majority of the cutting-edge forest sector research in Canada (Beaudoin *et al.* 2007; Frayret *et al.* 2007; Cid *et al.* 2008; Cambianghi *et al.* 2008; D'Amours *et al.* 2008, Feng *et al.* 2008; Forget *et al.* 2008a; Forget *et al.* 2008b; Gaudreault *et al.* 2009; D'Amours *et al.* 2010; D'Amours *et al.* 2011).

2.3 THE FOREST VALUE CHAIN

A fundamental problem in Canada's forest value chains is the information disconnect between forest and manufacturer planning departments. In attempting to quantitatively model the forest industry across the supply chain, forest sector issues become increasingly complex due to the increase in the goods and services provided by forests, and the linkages within the forest sector and with other sectors (Buongiorno 2003). There is also increasing pressure put on forest sectors in different countries due to an increasing globalized economy, stronger links between nations, trade liberalization and international treaties (Buongiorno 2003).

The Canadian forest industry has traditionally maintained a production focused industry model where manufacturers focused on producing. This model is changing to one that is customer oriented – also known as the pull-supply model – where the process starts with a market demand for specific forest product attributes. The pull-supply model supports diversified and value-added forest products that are linked directly to the attributes found at the tree level (Beaudoin *et al.* 2007; Groot and Pitt 2010). In the pull-supply mode the value chain needs to be mapped from consumer back to forest management and operations where particular wood fiber and characteristics are desired. This new type of industry design requires a heavy investment in decision support tools that can model and manage the entire value chain enterprise.

Value chains in the forest industry are related to supply chains, and both terms fall under the discipline of Operations Research (OR), which represents the study of optimal resource allocation to achieve specified objectives (Hillier *et al.* 1990; Sjostrom and Rask 2001; Beaudoin *et al.* 2007; D'Amours 2008; Jensen 2009). Similar to supply chain management the goal of value chain optimization is to maximize the value along all sections of the supply chain (Hillier *et al.* 1990; Sjostrom and Rask 2001; Beaudoin *et al.* 2007; D'Amours 2008). Using the word value instead of supply creates an impression that end product value is more important than merely tracking inventory across the supply chain. Although supply chain and value chain terms are often used synonymously it is important for the reader to note that an important shift occurs when we move from simply tracking supply to planning for maximum value. The forestry value chain is also defined as a linked web conceptual model of the entire forest industry, which can span several companies (Beaudoin *et al.* 2007; D'Amours 2008; Forget *et al.* 2008a). This is an important observation to make as companies traditionally have only shared supply chain information in-house. However, to optimize the entire supply chain, relevant data needs to flow freely between companies within the supply chain to truly maximize value for the entire group (Beaudoin *et al.* 2007; Frayret 2007; D'Amours 2008; Forget *et al.* 2008b). This shows that products and services flow from one contributor to another, and information exchanges keep the system in operation. Additional benefits of a value chain approach include the integration of decision-making based on wood/fibre yield, wood/fibre quality and production costs, and these pillars interact between the forest resources and extracting maximum value/benefit from them. Looking at the production process as a value chain can improve an enterprise through minimizing environmental impacts, procurement and production costs, and maximizing yield, quality, value and benefit to society (Pulkki 2004; Adams and Cavana 2009).

In order for a value chain to function, information must be shared freely across the entire enterprise (Frayret 2007; D'Amours 2008; a *et al.* 2008). The components in a value network include: raw material suppliers, the enterprise, extended enterprise and the final customers (Pulkki 2004; Beaudoin *et al.* 2007; Frayret 2007; D'Amours 2008; Forget *et al.* 2008a). There are several sub-groups within each of these components that add to the complexity of the value chain; information, currency and product variables travel through each component of the enterprise (Cambiaghi *et al.* 2008).

Once the network components have been established, the next stage is the development of an integrated business and logistics planning system (Frayret 2001; Sjostrom and Rask 2001; Forget *et al.* 2008b; D'Amours 2008). The integrated planning system is set up for the entire value chain in broad objectives, and then broken down by infrastructure, hierarchy and timescale (Frayret 2007; Gaudreault *et al.* 2009). There is no set pattern for breaking down the value creation network, but there are generally three layers of infrastructure: corporate, divisional and facility, as well as a timescale with multiple horizons (long, medium and short) depending on the objectives (Ronnqvist 2003; Frayret 2007; D'Amours *et al.* 2008; Gaudreault *et al.* 2009). For assessment strategies, the continual and accurate flow of information creates adaptation, assessment, improvement and feedback (Ronnqvist 2003; Beaudoin *et al.* 2007; Frayret 2007; D'Amours 2008; Forget *et al.* 2008b). For example, corporate strategic level planning includes business and supply chain strategies with a long-term timescale of five-plus years. Divisional planning involves the supply chain tactical objectives with a medium-term timescale of months. And operational planning is carried out by a facility under a minute-weeks horizon (Figure 1) (Ronnqvist 2003).

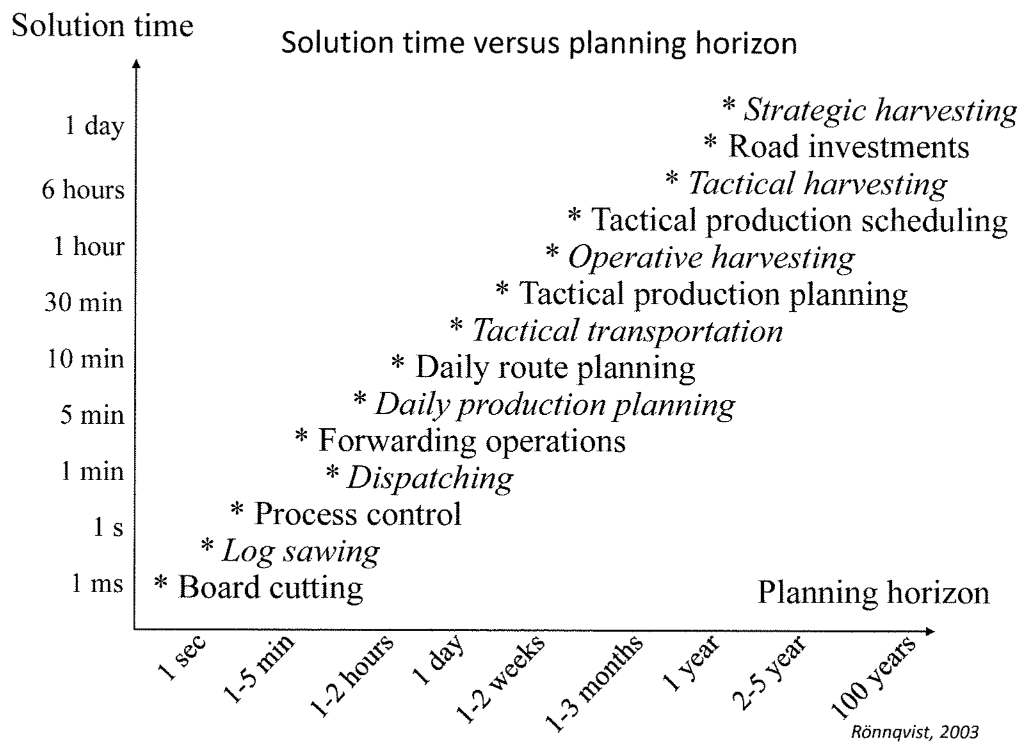


Figure 1. Solution time versus planning horizon for lumber mill supply chain optimization (Source: Ronnqvist 2003).

Information, product and currency travel between supply chain components at different timescales to satisfy objectives for each level of infrastructure. In order to optimize along the value chain, product inventory management is identified as having significant importance (Ronnqvist 2003; Beaudoin *et al.* 2007; Frayret 2007; D'Amours 2008; Forget *et al.* 2008b). The goal within the value creation network is to efficiently fulfill product demand, while simultaneously minimizing product inventory at each stage of the supply chain. Traditionally, forest industries in Canada have carried high inventories as a safety buffer against irregular wood supply. However, holding high inventories can carry severe negative impacts to the value network by incurring unneeded costs (Ronnqvist 2003; Beaudoin *et al.* 2007; Frayret 2007; D'Amours 2008; Forget *et al.* 2008a).

There is a threshold point in holding inventory where the cost to hold this inventory increases exponentially. Reasons for the rapid increase in cost include the facility cost of holding inventory, depreciation or decomposition of the product/resources, and market demands that can change rapidly resulting in wasted product/resources. Alternatively, when product supply levels fall below product demand value creation is lost (Montreuil *et al.* 2000; Sjostrom and Rask 2001). The failure to meet product demand at any stage along the supply chain creates gaps, as well as a literal breakdown of the supply chain (Montreuil *et al.* 2000; Sjostrom and Rask 2001). Due to value networks' dynamic nature and fluctuations in product demand, the optimal inventory level is constantly changing. Canadian research clusters are developing robust optimization models that can assist value creation networks by integrating information along the chain to minimize inventory and fulfill customer orders efficiently (Beaudoin *et al.* 2007a; Beaudoin *et al.* 2007b; Frayret *et al.* 2007; Cid *et al.* 2008; Cambianghi *et al.* 2008; D'Amours *et al.* 2008, Feng *et al.* 2008; Forget *et al.* 2008a; Gaudreault *et al.* 2009; D'Amours *et al.* 2010; D'Amours *et al.* 2011).

2.5 FOREST INVENTORY AND FOREST PRODUCT VALUE RECOVERY MODELING.

2.5.1 Product Recovery Modeling

Linking lumber products to an enhanced forest inventory has been receiving significant attention since the 1980's, but in Canada, this research focus of connecting enhanced forest resource inventory to simulated product recovery has only grown rapidly over the last decade (Pitt and Pineau 2009, Groot and Pitt 2010).

More recent modeling attempts are striving to predict product yields with a far greater degree of accuracy than seen previously. Many of these mo1dels are using state-of-the art technology such as airborne and terrestrial lidar, and advanced product recovery simulation tools (Beauregard *et al.* 2002; Malinen *et al.* 2003; Uusitalo and Isotalo 2005; Moberg and Nordmark 2006; Zhang *et al.* 2006; Murphy 2008). Modeling has been extended even further, striving to predict the inherent fibre properties using external tree scanning data (Li 2009; Groot and Pit 2010).

Detailed growth, recovery and fibre property simulators can provide a detailed assessment of lumber value yield. PipeQual and Optitek, for example, can simulate the internal 3D stem structure for use in product recovery modeling; used to simulate lumber yields, grades, and fibre properties (Grondin and Drouin 1998; Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a and 2007b; Kantola *et al.* 2008 and 2009). The simulators use input variables of tree age, dbh, height, crown ratio and diameter profiles (Grondin and Drouin 1998; Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a; Liu *et al.* 2007b; Kantola *et al.* 2008 and 2009).

Sawing simulation plays a critical role in forecasting lumber product recovery at the sawmill. Although computer controlled sawing simulators have been in existence since the 1970's (e.g., Halco Software Systems Ltd.), sawing recovery simulation has grown into a number of software developers with a variety of simulation platforms (e.g. Optitek, Sawsim, Halco and Comact) (Grondin and Drouin 1998; Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a; Liu *et al.* 2007b; Kantola *et al.* 2008 and 2009; Murphy *et al.* 2010a).

Log scanning technology is used in forecasting/predicting lumber product recovery from forest resources. Recent studies have shown the usefulness of using laser-scanning to connect external log features to lumber recovery variables which are then integrated into forest management software (Middleton *et al.* 2003; Murphy *et al.* 2010a). Log scanning technology has been used to develop forest management tools in Canada (e.g., the Forest Estate Enabling Software in British Columbia) designed for modeling and management of forest stands (Middleton *et al.* 2003). Research shows that sawing simulation software (e.g., Saw 2003) can be utilized to estimate lumber volume and grade recovery using pre-harvest data (Moberg and Nordmark 2006). Scanning research supports the observation that a volume-oriented forest management strategy does not necessarily lead to maximum product recovery and best return (Zhang *et al.* 1998). Forest stock dimensions, form and quality also need to be included in product recovery modeling (Zhang *et al.* 1998; Middleton *et al.* 2003; Moberg and Nordmark 2006).

Emerging product recovery models are striving to link external tree attributes to estimate internal wood quality attributes (Li 2009; Groot and Pit 2010). Mapping internal wood properties adds an additional layer of forest or tree detail supporting decisions on how to best match unique stand attributes to specific end products (Li 2009; Groot and Pit

2010). By integrating internal properties, these models also provide a basis for grading simulated products, lumber dimension, lumber value and defect allowance (Barbour and Kellogg 1990; Houllier *et al.* 1995; Wagner *et al.* 1996; Barbour *et al.* 1997; Ikonen *et al.* 2003).

Emerging product recovery research is using terrestrial laser scanning (TLS) to collect base information to model optimal bucking and product recovery (Carson *et al.* 2004; Murphy 2008). Lidar derived estimates of average stand value and log product yields were within 7% of actual hand-measured estimates (Murphy 2008). The TLS technology has been compared with cruised and manually measured log yield, and estimates of volume and value recovery are not significantly different (Murphy 2008). This research shows that TLS is a rapid and accurate method for detailed pre-harvest data acquisition and can be suitable for product recovery simulation (Murphy *et al.* 2008; Murphy *et al.* 2010b). Linking this type of product recovery modeling back to the forest stand for tactical and operational planning purposes has need of more research (Murphy *et al.* 2008; Murphy *et al.* 2010b).

Since the 1970's, researchers have worked to incorporate log value into the forest inventory or timber cruise. An important step in modeling value is simulated tree-bucking pre-harvest (Wagner 1996; Goulding 2000). Stem coding systems are used in product recovery modeling DSTs and support market-driven harvest planning based on enhanced forest inventory data such as dimension, form and quality of the standing trees (Wagner *et al.* 1996; Goulding 2000).

One such system is called MARVL, the Method for the Assessment of Recoverable Volume by Log-types (Deadman and Goulding 1979). MARVL produces a cruising framework that applies stem codes for each section of the stem. These stem codes contain quality and malformation information that can be entered into a computer model that bucks the stem into logs to forecast potential yield value from each stem (Deadman and Goulding 1979; Deadman 1990; Gordon and Lawrence 1995). A significant drawback of stem-coding cruising systems is that they are very time consuming and do not predict actual product breakdown (Gordon and Baker 2004). The dollar values optimized are those under ideal management conditions where there is a buyer for each log type produced (Gordon and Laurence 1995; Gordon and Baker 2004).

Additionally, the MARVL system was found to be too rigid for assigning stem parameters for more flexible recovery modeling (Gordon and Baker 2004). Researchers developed "Stem Description" as an alternative method to assign attributes along the stem (Gordon and Baker 2004). Stem description is found to be superior to stem coding in predicting volume by log grade per hectare (Gordon and Baker 2004). Problems with the stem codes were in efficiency and accuracy as there are over 18 codes for the cruiser to apply at once, when cruisers can usually only keep 9 codes in mind at once (Gordon and Lawrence 1995). A major advantage of the stem description method was shown in its ability to estimate log yields. In stem description, the data does not limit the range of potential simulated cutting strategies (Gordon and Baker 2004). Tree profiles collected provide a continuous range of description data along the stem. The stem description inventory method addresses the need for a more versatile coding system that log grade specifications and are increasingly more detailed while log prices are being more tightly linked to log quality (Deadman and Goulding 1979; Gordon and Lawrence 1995; Goulding *et al.* 1995; Gordon's and Baker 2004).

Dynamic tree bucking and product recovery models have been in existence since the 1970's. The trend is to take already established growth models and combine them with a user-interface geared towards forest practitioners in a stand-alone software platform (e.g., the Tree Value System). Input forest inventory data include basal area of the stand, DBH, tree height, heights to first live branch and first dead branch, and species-specific taper functions to determine optimal bucking strategies (Briggs 1989; Uusitalo and Kivinen 2000; Malinen *et al.* 2007). A more detailed pre-harvest measurement tool has been developed for predicting forest composition and generates an output of stand characteristics and tree size parameters (Uusitalo and Kivinen 2000). The tree list can be used to determine bucking parameters prior to harvest. Dynamic bucking simulation tools are useful in wood procurement and sawing production planning by simulating volume estimates for each product assortment (i.e., log, pulpwood, etc.) (Pnevmaticos *et al.* 1972; Briggs 1977; Briggs 1989; Uusitalo and Kivinen 2000; Malinen *et al.* 2007).

Additional research has found that local taper equations can be used to model individual tree stem profiles to simulate optimized bucking routines for specific log sizes and classes (Eng *et al.* 1986; Gobakken 2000; Zakrzewski *et al.* 2010). Additionally, the use of TLS has been used to model optimal bucking based on log specification (Pilkerton 2009; Murphy *et al.* 2010).

Comparison between the various types of data and analysis techniques to simulate optimal bucking and product recovery show that detailed stem profiling creates the most accurate results (Briggs 1989; Sachet *et al.* 1989; Malinen *et al.* 2007; Murphy *et al.* 2008). The research shows that taper functions can provide a good description of stem profiles, but cannot provide descriptions of sweep and branching within individual stems, thus further detailed stem assessments are required pre-harvest (Eng *et al.* 1986; Gobakken 2000; Zakrzewski *et al.* 2010). To support the forest industry management shift from volume to value there is the need for detailed stem profile information for product recovery simulation.

2.5.2 Important Tree and Log Variables

In modeling product recovery from tree-level and log-level inventory, it is important to measure external variables significant in predicting product recovery. When examining log and tree-level product recovery, there has been a large body of research modeling how external log features correlate with internal product recovery and value (Kellogg and Warren 1984; Steele 1984; Zhang and Gingras 1999; Beauregard *et al.* 2002; Wilhelmsson and Moberg 2004; Moberg and Nordmark 2006; Liu *et al.* 2007a, Liu *et al.* 2007b). Important tree variables include: diameter at various heights along stem (stem profile), diameter at breast height, height of tree, height to lowest dead branch, height to lowest live branch, crown length (from lowest live branch to tree top), taper, log or stem shape (sweep, crook, eccentricity), knot frequency and size and internode index (Kellogg and Warren 1984; Steele 1984; Zhang and Gingras 1999; Zhang *et al.* 2001; Beauregard *et al.* 2002; Wilhelmsson and Moberg 2004; Moberg and Nordmark 2006; Liu *et al.* 2007a, Liu *et al.* 2007b).

Many trends in modeling lumber volume and value recovery are similar among researchers. For example, tree DBH had the greatest effect on modeling tree-level product value, followed by tree height and tree taper (Beauregard *et al.* 2002; Wilhelmsson and Moberg 2004; Moberg and Nordmark 2006). Increases in taper decreased the lumber value recovery. For intensively managed stands, the modeling results showed that pruned logs graded significantly higher than un-pruned logs (Steele 1984; Beauregard *et al.*

2002; Ruel *et al.* 2010). The results also showed that for un-pruned logs, the longer the internodal length, a higher yield value was realized; demonstrating the importance of including branch diameter, frequency and distribution along logs and stems when modeling lumber value recovery (Steele 1984; Beauregard *et al.* 2002; Ruel *et al.* 2010).

Studies cataloguing the variables found to be significant in lumber recovery in jack pine found DBH, total tree height, height to first live branch and first dead branch, knot size, branch diameter, crown length/ratio, branch index and basal area of the stand had the greatest influence on value recovery (Steele 1984; Zhang *et al.* 2001; Zhang and Tong 2005; Zhang *et al.* 2006; Ruel *et al.* 2010). Other stem characteristics such as sweep, crook and eccentricity are statistically significant but do not impact lumber value recovery a great deal (Steele 1984; Ruel *et al.* 2010).

Other studies have simply used DBH, height and taper to predict lumber output and lumber grade yield in a standard- and random-length sawmill (Zhang and Tong 2005; Zhang *et al.* 2006; Liu 2007a; Liu 2007b). Multiple regressions were applied and explained up to 86% of observed variation in lumber grade yield and up to 92% in lumber value (Zhang *et al.* 2001; Zhang and Tong 2005; Zhang *et al.* 2006; Liu 2007a; Liu 2007b). The literature suggests for a lower-cost estimate of lumber grade yield, that linear regression models using DBH alone or combined with tree height would be effective in lumber grade yield and lumber value prediction. (Zhang *et al.* 2001; Zhang and Tong 2005; Zhang *et al.* 2006; Ruel *et al.* 2010).

The majority of lumber recovery models created for Canadian sawmills are general and reflect what is termed as a "normal" sawmill (Zhang *et al.* 2001; Zhang and Tong 2005; Zhang *et al.* 2006; Liu 2007a; Liu 2007b). However, most sawmills are unique, thus general models may not be appropriate for individual sawmills striving to model lumber value recovery from forest inventory data. Efforts should be made to calibrate locally valid models for each sawmill's lumber value creation process. With mill-specific models supporting forest value chain DSTs, management decisions can be made in the context of mill-specific product recovery to achieve specific production objectives.

2.7 PRINCIPAL COMPONENT AND REGRESSION ANALYSIS USED IN LUMBER RECOVERY STUDIES

Principal Component Analysis (PCA) has been used in several lumber value recovery studies to isolate components from a list of measured variables (Bharati *et al.* 2003; Chiorescu and Gronlund 2004; Flodin *et al.* 2008; Jones and Emms 2010). PCA is employed because most of the variables measured at the log- or tree-level have collinearity (Eriksson *et al.* 1999: Roos *et al.* 2000).

The use of PCA and PC regression scores in modeling can reduce variables that are collinear into factors, thus reducing the level of complexity in the model (Eriksson *et al.* 1999: Roos *et al.* 2000). In general, PCA is a well-established method for making sense of large data sets of biological material such as trees, where there is collinearity and noise (Eriksson *et al.* 1999: Roos *et al.* 2000).

PCA has been used in lumber recovery studies. For example, PCA has been employed in log tracking modeling in Sweden to track logs from log sorts to sawlines, in order to create an advanced raw material flow control where each log has a unique signature based on its laser-scanned external log characteristics (Uusijarvi 2000; Chiorescu and Gronlund 2004; Flodin *et al.* 2008). Other studies have used PCA to reduce log scanning data into Principal Components (PC) to track logs through a production chain by means of multivariate PCs and Most-Similar-Neighbour (MSN) analysis (Bharati and MacGregor 2003). In log tracking, the PCA signature was superior to using any single measured log parameter at one time. PCA has also been used to research softwood lumber image analysis, by running a multi-way PCA to decompose the acquired three-dimension lumber images into a two-dimensional principal component space (Bharati and MacGregor 2003). These components were used in predicting lumber recovery features in the 2D analysis space. PCA showed strong correlation between externally measured variables (i.e., branch size variables, the number of branches and whorls, mean internode length and log acoustic velocity) and internal properties (i.e. green density and heartwood content) (Bharati and MacGregor 2003; Jones and Emms 2010).

PCA has also been used at the tree-level in integrating lumber value recovery analysis to predict lumber value (Roos *et al.* 2000; Liu *et al.* 2007a; Liu *et al.* 2007b). PCA has been employed to describe the interrelationship of tree-level variables (i.e., DBH, height and taper). Due to the low number of input variables, PCs were not used later in the lumber recovery modeling; instead the original, measured variables were used in analyses (Roos *et al.* 2000; Liu *et al.* 2007a; Liu *et al.* 2007b).

The majority of lumber and tree recovery models utilize log and tree characteristics such a log size, geometry and quality (Steele 1984; Middleton *et al.* 1989; Liu *et al.* 1989; Shi *et al.* 1990; Wagner and Taylor 1993; Roos *et al.* 2000; Zhang and Tong 2005; Liu *et al.* 2007a; Liu *et al.* 2007b). Modelers have used regression methods along with optimized programming to relate these tree characteristics to lumber recovery variables. Generally, measured variable are used in regression analysis to examine how tree and stand-level characteristics can predict lumber value, lumber grades and internal properties based on stand and tree characteristics in conifers (Steele 1984; Middleton *et al.* 1989; Liu *et al.* 1989; Shi *et al.* 1990; Wagner and Taylor 1993; Roos *et al.* 2000; Zhang and Tong 2005; Liu *et al.* 2007a; Liu *et al.* 2007b).

In conclusion the research shows that PCA and regression analyses techniques are commonly used in lumber recovery studies. Also, PCA has been shown as an effective method to reduce a large number of measured tree variables into a few components. Linear and nonlinear models have also been described as an effective method for modeling product recovery parameters. For this reason PCA and regression techniques were applied in the modeling for this study.

3. METHODOLOGY

3.1 STUDY LOCATIONS

Sites used in this study were located in the Boreal forest region in northern Ontario north of Thunder Bay (Figure 2). We collaborated with Resolute Forest Products (previously Abitibi-Bowater) and their harvesting contractor Marcri Logging to carry out the operational planning and harvesting for the study; including harvesting, tracking of tree-lengths and chips from forest stands to the sawmill and pulp and paper mill in Thunder Bay, and sawmilling of the jack pine tree-lengths. The three study sites were selected from the summer and fall harvest plan of 2010 to cover a range of stem qualities for jack pine. Table 1 describes the three study site conditions. The three sites were located on similar sites ranging from moderately fresh and sandy to fresh coarse loamy. All sites were highly productive. The site classes are based on Plonski's single-species, normal yield tables (Plonski 1960) and therefore, should be applied to mixedwoods with caution.

Table 1. Stand characteristics by block.

Block	383	559	568
Ecosite Classifcation (ELC v.2.0 2009)	B049Tt/Tl	B034Tt/Tl	B050Tt/Tl
Soils	Fresh, Sandy	Dry-Fresh, Sandy	Fresh, Coarse Loamy
Site Class (Plonski)	1	1-2	1
Species Mix	$Pj_7Sb_2Bf_1$	$Pj_7Sb_2Po_1$	$Pj_4Po_3Sb_1Bf_1Bw_1$
Pj, dominant age (years)	115-125	90-100	80-100
Density (SPH)	872	828	658
All spp. gross merch vol (m^3/ha)	255.7	217.5	282.1
All spp. net merch vol (m^3/ha)	235.6	202.7	256.8
Pj, gross merch vol (m^3/ha)	186.7	140.0	106.9
Pj, net merch vol (m^3/ha)	169.9	133.3	101.7
Pj, tree-length vol (m^3/ha)	135.9	106.6	81.4
Pj, average net merch tree size (m^3)	0.7	0.4	0.5
Pj, average defect severity	Med	Low	High

The sites were harvested using a conventional mechanized full-tree system. Sawlog wood was transported to the sawmill in tree-length form. The undersized and non-lumber species were chipped using the chain flail delimber-debarker chipping system. Approximately 500 m^3 of jack pine tree-lengths from each site were delivered to the Resolute Forest Products sawmill in Thunder Bay. The 10 truckloads from each site were harvested the week before and delivered on the Friday night before each mill run. On three separate Saturday mornings (6:00 a.m. – 12:00 p.m.), the jack pine volumes from each site were processed independently through the sawmill after all logs, lumber and residue were purged from the saw lines, and all computers scanning information and production data were zeroed. Information was tracked at each stage of the wood procurement and conversion process up to rough green lumber production. The datasets collected for each site included a pre-harvest inventory, tree-length taper and quality profiles, site and mill recovery volumes and products, residual biomass and standing tree volume, and internal fibre properties of sample jack pine trees.

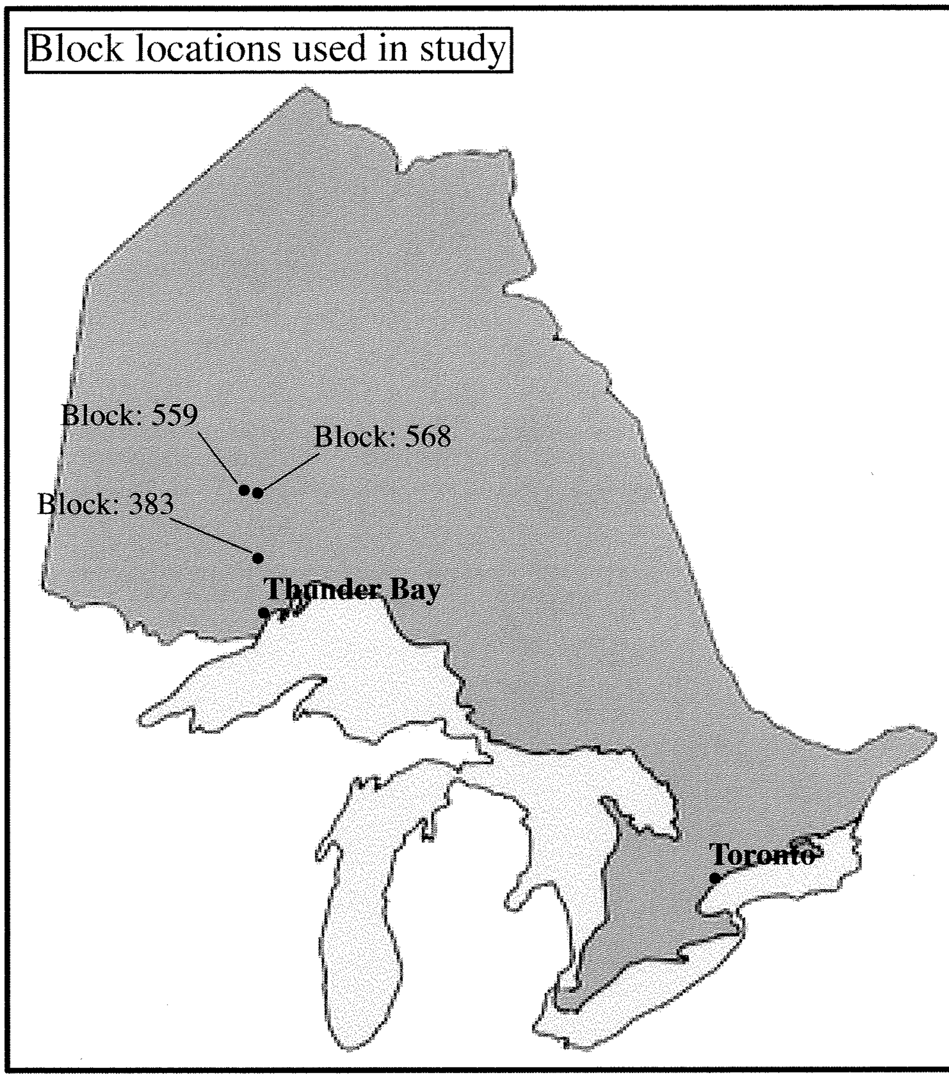


Figure 2. Map showing locations of three study sites in northern Ontario (Source: Ministry of Transportation Ontario).

3.1.1 Study Site 1

Block 383 was the first site inventoried and was processed at the sawmill on June 12, 2010. The terrain was gentle with a slight SW slope (2 - 6%). The Ecosite Classification was B049Tt/Tl (OMNR 2009) and had a Forest Ecosystem Classification of V17 (OMNR 1997) with fresh, sandy mineral soils. A picture showing the average stand conditions for block 383 is seen in Fig 3. The site was class 1 (Plonski 1960), with stocking approximately 658 stems per hectare (SPH). Approximately 3.7 ha was harvested from the block to run a 500 m^3 sample through the sawmill.

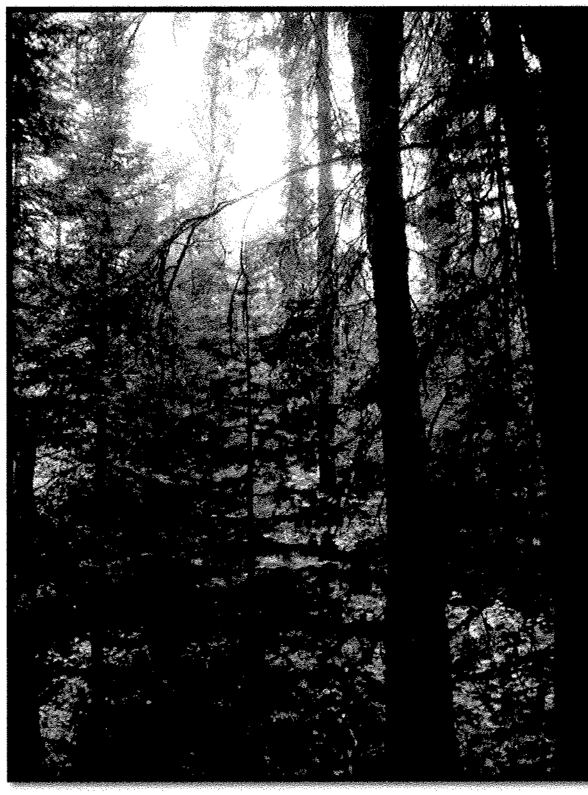


Figure 3. A picture of the typical stand conditions found within block.

The study site of block 383 had a gross merchantable volume of 255 m^3/ha with a volume composition of 73 % jack pine, 21 % spruce (*Picea spp.*), 4 % balsam fir (*Abies balsamea* (L.)(Mill.)), and 2 % trembling aspen (*Populus tremuloides* (Michx.)) (Table 2). The net merchantable jack pine volume was 169.9 m^3/ha (Table 3). Form defects were found on 49 % of jack pine stems. Defects were moderate in severity. A high amount of scaring and large diameter branches were present in the jack pine. Spruce in the stand was of moderate quality with high branchiness. Balsam fir was of moderate quality with scarring present on 25 % of larger diameter stems. Trembling aspen was of moderate to poor quality with a high concentration of conk on over 50 % of stems.

Table 2. Gross merchantable volume by species and diameter class for block 383.

Block: 383	Gross Merch Volume (m^3/ha)							
# Plots = 12	Diameter Class							
Species	**0 - 9.9**	**10 - 13.9**	**14 - 17.9**	**18 - 21.9**	**22 - 25.9**	**26 - 29.9**	**30+**	**Total**
Bf	0.0	3.2	2.2	1.9	2.4	0.0	0.0	**9.7**
Bw	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Pj	0.0	0.0	0.0	7.0	6.8	9.5	163.4	**186.7**
Pt	0.0	0.0	0.0	0.0	0.0	3.9	0.0	**3.9**
Sb	0.0	5.2	10.8	8.7	11.9	4.1	14.7	**55.4**
Sw	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Total	**0.0**	**8.4**	**13.0**	**17.6**	**21.1**	**17.6**	**178.1**	**255.7**

Table 3. Net merchantable volume by species and diameter class for block 383.

Block: 383	Net Merch Volume (m^3/ha)							
# Plots = 12	Diameter Class							
Species	**0 - 9.9**	**10 - 13.9**	**14 - 17.9**	**18 - 21.9**	**22 - 25.9**	**26 - 29.9**	**30+**	**Total**
Bf	0.0	3.1	2.1	1.8	2.3	0.0	0.0	**9.3**
Bw	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Pj	0.0	0.0	0.0	6.4	6.2	8.7	148.7	**169.9**
Pt	0.0	0.0	0.0	0.0	0.0	3.1	0.0	**3.1**
Sb	0.0	5.0	10.4	8.4	11.5	4.0	14.1	**53.3**
Sw	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Total	**0.0**	**8.1**	**12.5**	**16.5**	**20.0**	**15.7**	**162.8**	**235.6**

3.1.2 Study Site 2

Block 559 was the second site inventoried and was processed at the sawmill on October 2, 2010. The terrain was gentle- rolling with SW slope (1-12%). The Ecosite Classification was B034Tt/Tl (OMNR 2009) and had a Forest Ecosystem Classification of V32 (OMNR 1997) with dry-fresh, sandy mineral soils. A picture showing the average stand conditions for block 559 is seen in Figure 4. The site was class 1-2 (Plonski 1960) with stocking approximately 828 SPH. The area harvested from this block to deliver 500 m^3 to the sawmill was 4.7 ha.

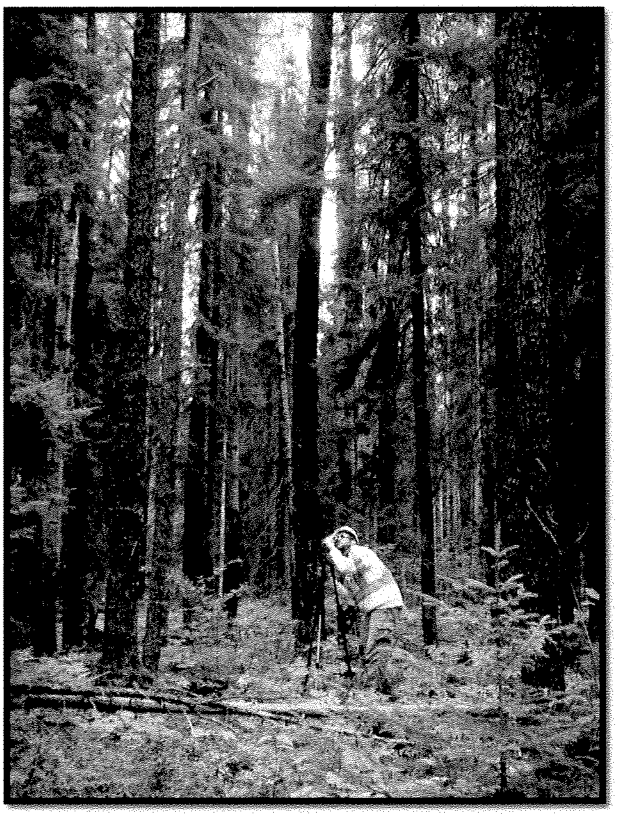


Figure 4. A picture of the typical stand conditions found within block.

The study section of block 559 had a gross merchantable volume of 217 m^3/ha with a volume composition of 64 % jack pine, 23 % black spruce (*Picea mariana* (Mill.)B.S.P.), 12 % trembling aspen, and 1 % balsam fir (Table 5). The net merchantable jack pine volume was 133.3 m^3/ha (Table 6). The jack pine was of moderate to good quality with form defects on 30 % of stems with low to moderate severity. The black spruce was of moderate to good quality with minor form defects. The aspen was of moderate quality with conk defect found on 25 % of stems. The balsam fir was of moderate quality, small diameter and in the understory.

Table 5. Gross merchantable volume by species and diameter class for block 559.

Block: 559	**Gross Merch Volume (m^3/ha)**							
# Plots = 10	**Diameter Class**							
Species	**0 - 9.9**	**10 - 13.9**	**14 - 17.9**	**18 - 21.9**	**22 - 25.9**	**26 - 29.9**	**30+**	**Total**
Bf	0.0	2.1	0.0	0.0	0.0	0.0	0.0	**2.1**
Bw	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Pj	0.0	0.0	2.0	12.4	50.1	43.2	32.2	**140.0**
Pt	0.0	0.0	0.0	0.0	3.3	0.0	23.3	**26.7**
Sb	0.0	4.1	9.0	18.0	8.7	9.0	0.0	**48.8**
Sw	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Total	**0.0**	**6.2**	**11.0**	**30.4**	**62.1**	**52.2**	**55.6**	**217.5**

Table 6. Net merchantable volume by species and diameter class for block 559.

Block: 559	**Net Merch Volume (m^3/ha)**							
# Plots = 10	**Diameter Class**							
Species	**0 - 9.9**	**10 - 13.9**	**14 - 17.9**	**18 - 21.9**	**22 - 25.9**	**26 - 29.9**	**30+**	**Total**
Bf	0.0	2.0	0.0	0.0	0.0	0.0	0.0	**2.0**
Bw	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Pj	0.0	0.0	1.9	11.8	47.7	41.2	30.7	**133.3**
Pt	0.0	0.0	0.0	0.0	2.5	0.0	17.9	**20.5**
Sb	0.0	4.0	8.7	17.3	8.4	8.7	0.0	**46.9**
Sw	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Total	**0.0**	**6.0**	**10.6**	**29.1**	**58.6**	**49.8**	**48.6**	**202.7**

3.1.3 Study Site 3

Block 568 was the third site inventoried and processed at the sawmill on October 16, 2010. The terrain was gentle with a slight W slope (0 - 4%). The Ecosite Classification was B050Tt/Tl (OMNR 2009) and had a Forest Ecosystem Classification of V11 (OMNR 1997) with fresh, course loamy mineral soils. A picture showing the average stand conditions for block 559 is seen in Fig 5. The site was class 1 (Plonski 1960) with stocking approximately 872 SPH. The area harvested from the block to deliver 500 m^3 to the sawmill was 6.5 ha.

Figure 5. A picture of the stand conditions found within block 568.

The study section of block 568 had a gross merchantable volume of 282 m^3/ha with a volume composition of 38 % jack pine, 35 % trembling aspen, 11 % black spruce, 8 percent balsam fir, 7 % white birch (*Betula papyrifera* Marsh.) and 1 % white spruce (*Picea glauca* (Moench) Voss) (Table 7). The net merchantable jack pine volume was 96.2 m^3/ha (Table 8). The jack pine was of moderate to low quality with high form defects and large branches. The average lowest dead branch was lowest of the three sites (3.5 m). Form defects were found on 75 % of jack pine stems and were moderate to high in severity. The jack pine also had a high amount of scarring. The trembling aspen was of moderate quality with a high concentration of conk on large diameter stems. Black and white spruce were moderate quality with high branchiness. The balsam fir was of moderate quality and mainly found in the understory in thick patches. The white birch was of moderate to poor quality with high form defect and primarily found in the understory.

Table 7. Gross merchantable volume by species and diameter class for block 568.

Block: 568	**Gross Merch Volume (m^3/ha)**							
# Plots = 20	**Diameter Class**							
Species	**0 - 9.9**	**10 - 13.9**	**14 - 17.9**	**18 - 21.9**	**22 - 25.9**	**26 - 29.9**	**30+**	**Total**
Bf	0.0	2.6	5.0	9.3	4.7	1.6	0.0	**23.1**
Bw	0.0	3.7	4.3	6.0	4.5	0.0	0.0	**18.5**
Pj	0.0	0.0	0.0	8.4	17.9	22.2	58.3	**106.9**
Pt	0.0	0.0	1.5	12.0	9.9	14.4	60.1	**97.9**
Sb	0.0	2.3	4.0	5.8	7.8	3.3	8.9	**32.2**
Sw	0.0	0.0	0.0	1.6	0.0	0.0	1.8	**3.4**
Total	**0.0**	**8.6**	**14.9**	**43.1**	**44.8**	**41.6**	**129.2**	**282.1**

Table 8. Net merchantable volume by species and diameter class for block 568.

Block: 568	**Net Merch Volume (m^3/ha)**							
# Plots = 20	**Diameter Class**							
Species	**0 - 9.9**	**10 - 13.9**	**14 - 17.9**	**18 - 21.9**	**22 - 25.9**	**26 - 29.9**	**30+**	**Total**
Bf	0.0	2.4	4.6	8.7	4.3	1.5	0.0	**21.5**
Bw	0.0	3.6	4.3	5.8	4.4	0.0	0.0	**18.1**
Pj	0.0	0.0	0.0	7.6	16.1	20.0	52.5	**96.2**
Pt	0.0	0.0	1.3	9.9	8.2	11.9	49.8	**81.2**
Sb	0.0	2.2	3.9	5.6	7.5	3.2	8.6	**31.0**
Sw	0.0	0.0	0.0	1.5	0.0	0.0	1.8	**3.3**
Total	**0.0**	**8.2**	**14.0**	**39.1**	**40.6**	**36.7**	**112.7**	**251.3**

3.2 FIELD MEASUREMENTS

Table 9 presents the datasets measured across the value chain studied. Stages included pre-harvest timber cruise, tree-length profiling, mill processing, residual biomass inventory and internal properties testing. It is important to note that the analysis section uses timber cruise and sawmill recovery datasets only and the remaining datasets are described as a point of reference for future research.

Table 9. Description of data variables collected at each stage of the study.

Stage 1: Timber Cruise	**Stage 2: Tree-length profiling**	**Stage 3: Mill Processing**	**Stage 4: Residual Biomass**	**Stage 5: Internal Properties**
Objectives: 1. Establish georeferenced prism cruise plots to record stand volume and quality parameters.	**Objectives:** 1. Record stem taper and quality profiles.	**Objectives:** 1. Track volume (chips and sawlogs) by species. 2. Pull detailed lumber recovery files from mill scanners.	**Objectives:** 1. Record residual biomass and standing tree volume.	**Objectives:** 1. Determine strength properties in jack pine sample trees at each site.
Data Collected: 1. Plot Profile • Aspect • Slope • Ecosite • Crown closure • Dominant species age • Azimuth and distance to each tree from plot center 2. Tree height at: • Top • 1st dead branch • 1st live branch • Top merchantable height 3. Diameter at: • DBH • 2.5m increments 4. Visual quality defect: • Crook • Sweep • Snapped top • Fork • Frost crack • Mechanical damage • Scar • Snag 5. Defect severity: • Scale 1-5 (1 is barely noticeable, 5 is non-merchantable) 6. Defect range: • Start and stop position on tree	**Data Collected** 2. Length-diameter measures: • Diameter at: base, DBH, 2.5m increments, top 3. Branch measures: • Max branch diameter at each 2.5m increment 4. Visual quality defect: • Heart rot • Ridge • Crook • Sweep • Snapped top • Fork • Frost crack • Mechanical. damage • Scar • Snag 5. Defect severity: • Scale 1-5 (1 is barely noticeable, 5 is non-merchantable) 6. Defect range: • Start and stop position on tree-length	**Data Collected:** 1. Tree-length and chip • Weigh-scale reports from sawmill and pulpmill 2. Tree length to log scale: • Tree length • Size of logs recovered • Quality of logs recovered 3. Optimization sawing scanners: • Date & time • Log diameter small • Log diameter large • Log length • Log volume • Sweep • Taper • Log grade • Lumber of piecies in log • Individual piece volume, grade • Lumber recovery value for log • Primary lumber recovery 4. Waste and Residue data: • Volume (end pieces, chips, sawdust)	**Data Collected:** 1. Ground biomass transect method: • Length • Diameter • Species • Type • Quality 2. Residual standing trees using BAF 2 prism method: • DBH • Height • Species • Quality	**Data Collected:** 1. Full log measurements of sample trees: • Same measurements as stage 2 2. Bolts (1m) cut from 0, 25, 50, 75, and 100 percent of merchantable stem (top diameter 10cm outside bark: • MOE • MOR • Compression • Density 3. Disks cut at 2.5m intervals for ring growth and density profiles pith to bark: • Early wood/Latewood diameters • X-ray densitometer profile from pith to bark

3.2.1 Pre-harvest Timber Cruise

Stage 1 data collection consisted of a pre-harvest timber cruise. The method used was a modified random point design with random plot centres as described by Avery and Burkhartt (2001). A prism BAF of 4 was used due to the well-stocked nature of the stands that yielded an average count of 11 trees per plot. A sampling intensity of 2 plots per hectare was taken. There were time restraints on the window available to conduct the timber cruises due to harvest scheduling. Therefore, a more intense sampling intensity was employed to ensure that the coefficients of variation of species volume by stand were acceptable. It would have been preferred to use a double sample or 3P sampling technique as outlined by Avery and Burkhartt (2001), but due to the short timeline on harvest scheduling and distance of study sites from town, it was only possible to conduct an intensive single-pass point cruise.

In addition to traditional variables measured in a point cruise, quality information was collected for jack pine stems as outlined in Table 9 to include quality, form and branchiness characteristics. The variables chosen for collection in the pre-harvest timber cruise are based on the body of research presented in the literature. The stem characteristics data collection technique was based on the surveying technique presented by Gordon and Baker (2004), where stem description information is not confined to specific log grades and can be quantified later for simulating a variety of bucking strategies. Figure 6 illustrates how the cruiser views the tree when categorizing and recording tree profile and quality characteristics during the timber inventory.

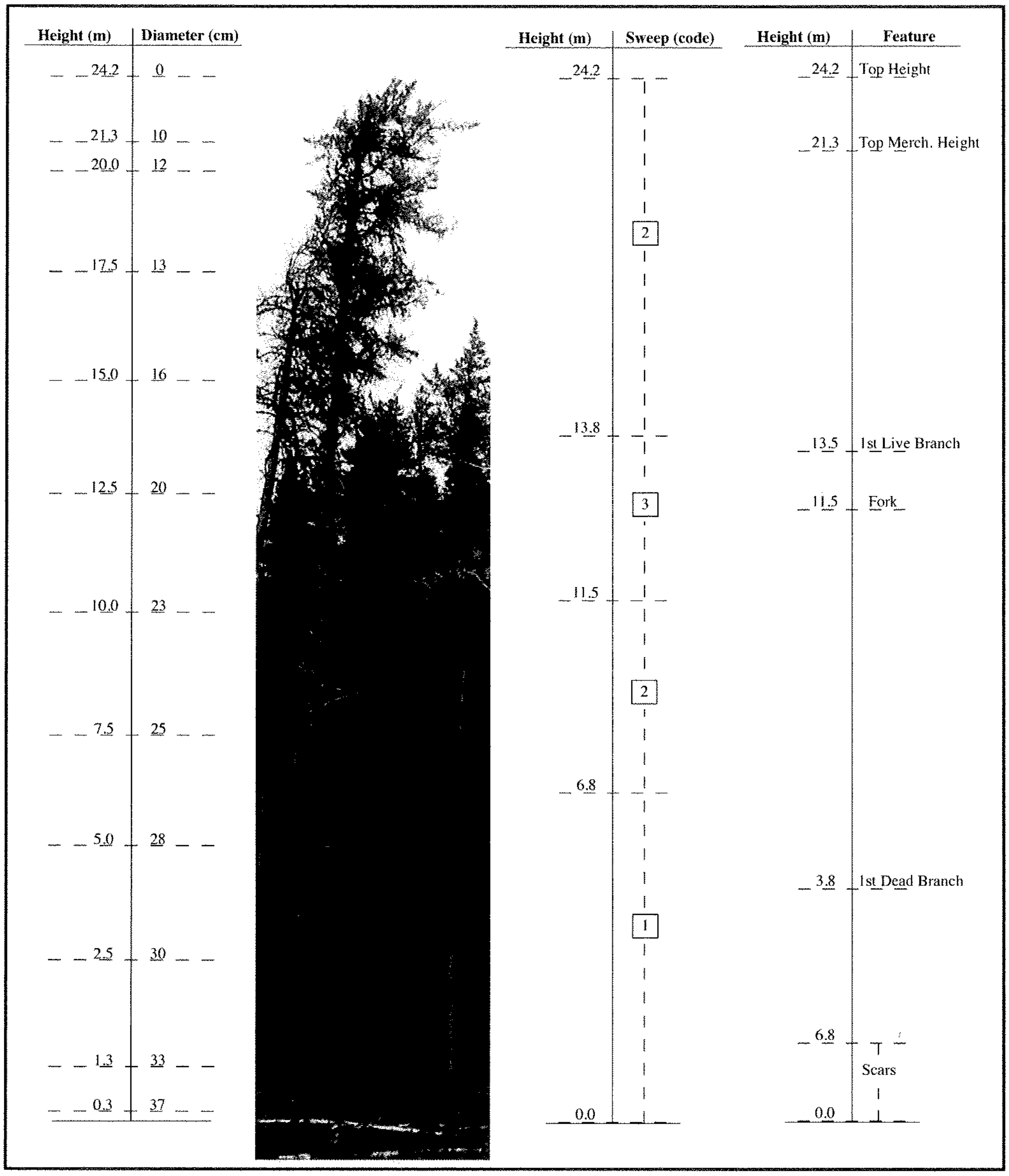


Figure 6. Tree profiles created during the timber cruise for merchantable jack pine stems describing a diameter, sweep and feature profile.

Stem profiles (height/diameters) were produced using the Laser Tech Criterion RD1000, an electronic version of a relascope. Horizontal distances were calculated using the Hagglov Vertex and input into the Criterion to automatically calculate the heights. On the Criterion measuring screen, manually adjusted bars are displayed to compute diameters. Field data for stages 1 and 2 were collected in a mobile PC unit with a custom-built form using Gerema Mobile software. The Gerema mobile software is an extension of ArcGIS mobile software in which Central Computer Services CCS has built a new application based off of the ArcGIS background for custom forest surveying needs.

Stage 1 tree diameter measures at DBH were taken using a diameter tape; distances and azimuth to plot centre were done using the Hagglov Vertex and a Silva Compass. Data was digitally recorded from the Criterion RD1000 (Figure 7) and Vertex IV into the Gerema Mobile handheld units.

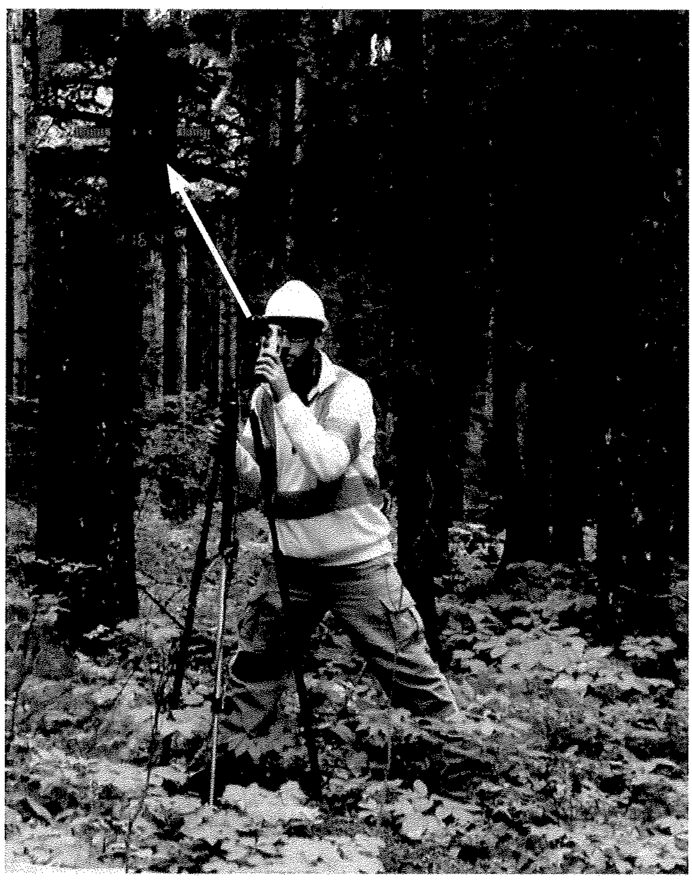


Figure 7. Criterion RD1000 electronic relascope being used to record stem profile data.

Diameter profiles of trees were measured using an electronic relascope, following a methodology from Gordon and Baker (2004), which was based off of original work by Deadman and Goulding (1979) using the Spiegal Relascope. In this method, characteristics (size, form, quality) are measured along the stem to produce a stem profile. The stem profile is used in future modeling to simulate the product recovery process (bucking and sawing simulation) and assign a product recovery value. In this study, we used the LaserTech Criterion RD1000 for measuring heights and diameters; count trees,and horizontal distances were measured using the Vertex IV.

3.2.2 Tree-length Profiling

Tree-length measurements (Stage 2 data collection) at roadside were taken to provide a snapshot of jack pine tree-lengths taken to the sawmill. The variables measured during the tree-length profiling are listed in Table 9. The basic tree-length measurements are the same as in the pre-harvest timber cruise; providing a full stem description including a diameter profile, sweep profile, top merchantable height and defect ranges. The profiling also measures maximum knot diameter for each 2.5 m section for each tree-length measured.

Stage 2 data was collected using calipers, a 30 m measuring tape, and the Gerema Mobile data-recording device with custom built data forms for this study. Stem profiles with quality and form characteristics were recorded (Table 9). Also, the butt ends of the logs were marked with blue paint for identification at the sawmill (Figure 8).

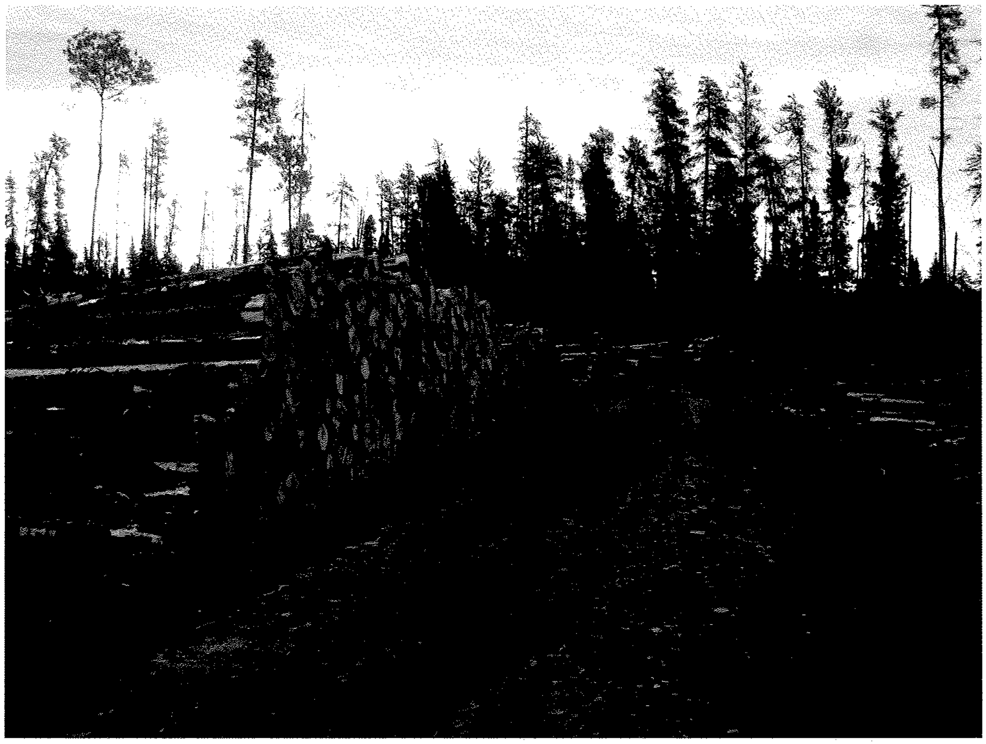


Figure 8. Tree-length stacked at roadside marked for study.

3.2.3 Mill Processing

Mill recovery data (Stage 3 dataset) were recorded at the Resolute Forest Products sawmill and pulpmill in Thunder Bay. The data were collected at four stages of the mill processing operation: tree-length weigh scale at sawmill, chip truck weigh scale at pulpmill, sawmill log and lumber optimizing scanner, and chips and sawdust data from the lumber recovery process weigh scaled at pulpmill. The purpose of collecting these datasets was to provide both an overall recovery of product from the study sites (tree-length, forest chips, fibre, sawmill chips and sawdust), and a detailed understanding of the lumber value and volume recovery. Quality and volume specifications were recorded for tree-lengths and chips delivered to both the sawmill and pulpmill. Residues from the sawing process, both chips and sawdust, were recorded after each site was run through the sawmill.

Sawmill measurements were conducted at the Resolute Forest Products random lengths sawmill in Thunder Bay, Ontario. The milling process moves the raw wood material through the wood-to-lumber conversion process (Figure 9) and records scanner information of size, recovery and waste parameters (Table 9**)**. Specifically for this study, we examined the Comact sawline scanner optimizers, which create a 3-D image of the log and run through cutting patterns until optimal log value recovery is reached. The log is then sawed and the lumber is send through trimmers, reclaimers, and then sorted by size and moisture content for kiln drying. After drying, the lumber runs through the planer and then grading. Due to the set up at mill, we were only able to track lumber from each site up to the rough-green lumber sort. Detailed variables measured during sawmilling process are listed in Table 9.

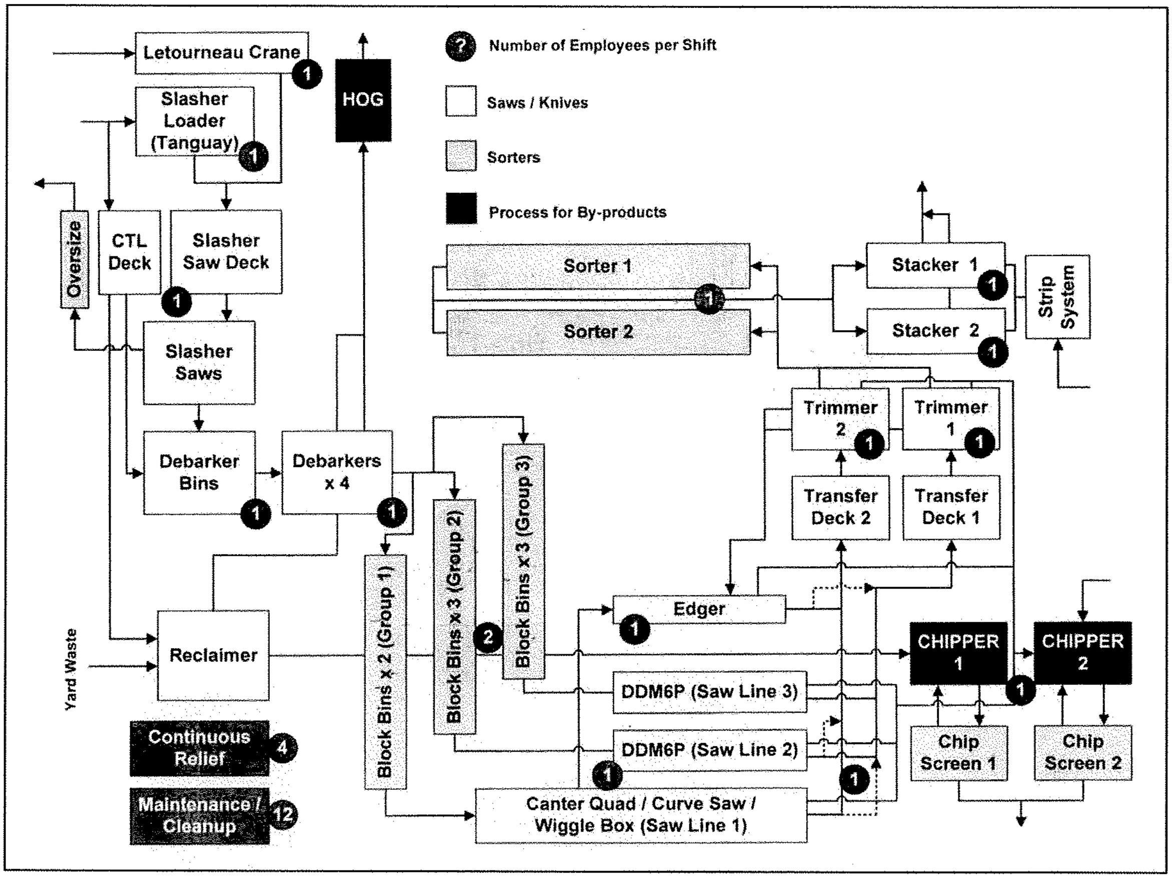


Figure 9. Sawmill wood processing flowchart (Leclerc 2011).

3.4 DATA ANALYSIS

Data was compiled into Microsoft Excel sheets through transcription, data transfer from Gerema Mobile data logger, and CSV files produced from the Comact scanner/optimizers. The statistical analyses were performed using IBM software PASW18. Figure 10 shows a flowchart outlining the data analyses sequence conducted in the study.

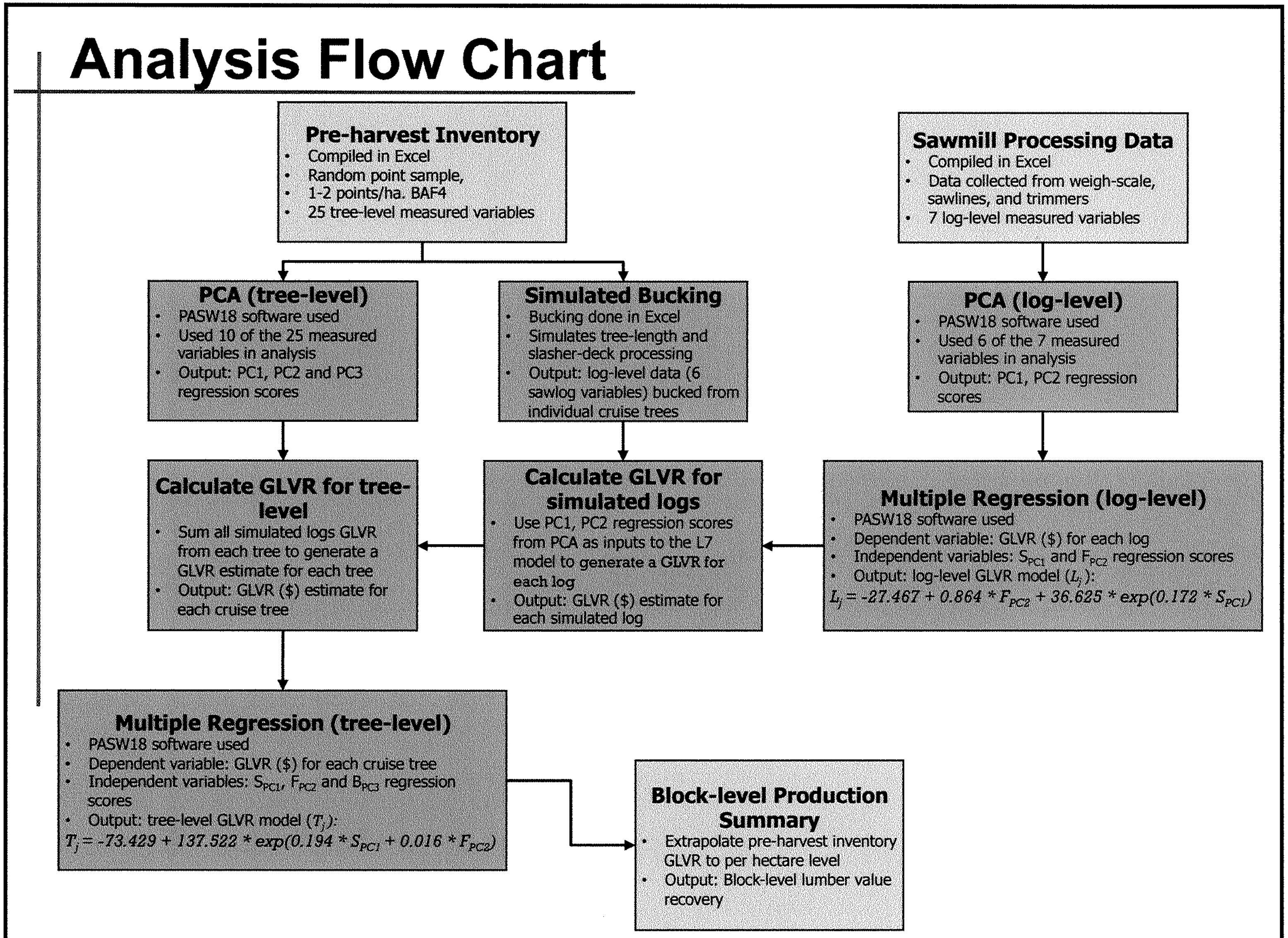


Figure 10. Flowchart describing the data analysis sequence conducted in the thesis.

3.4.1 Tree Bucking Simulation

A tree stem-bucking simulator was created in Microsoft Excel to simulate the slasher deck process at the sawmill using the pre-harvest inventory. Figure 11 shows the slasher deck at the sawmill that the tree bucking simulation is based upon. The bucking simulator extracted the same log-level variables (large-end diameter, small-end diameter, length, taper, volume and sweep) used in log-level PCA and regression analysis.

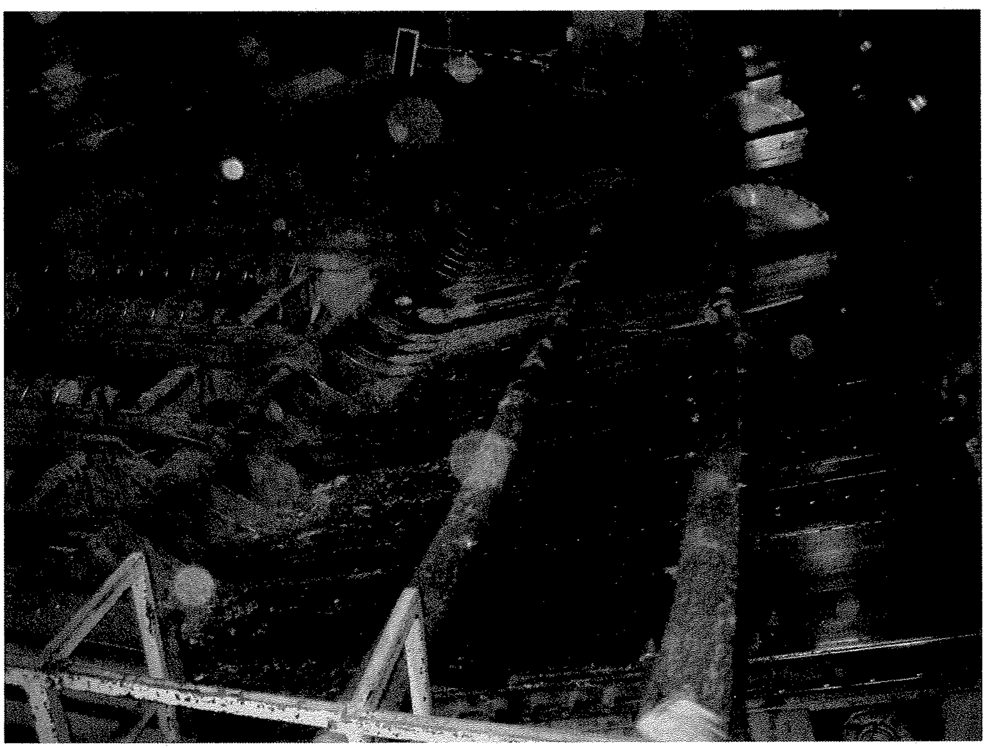


Figure 11. Picture of the slasher deck at sawmill bucking tree-lengths into logs.

The simulated bucking process was used to simulate the logs that would be recovered from each jack pine tree in the timber cruise. For this purpose, Excel spreadsheets were used to transform the pre-harvest timber cruise data into a form that can be interpreted for bucking patterns set up at the sawmill. The random lengths sawmill bucking process is not optimized; rather it is a set cutting pattern of 3.17 m, and all subsequent lengths of 2.51 m. Log diameters for each simulated log were calculated by interpolating diameters from the 2.5 m interval diameters taken during timber cruise with the electronic relascope. Stem characteristics (e.g. sweep, defects, etc.) were copied to each log bucked from stem in those sections (see figure 6 for stem details). All of the semi-automated bucking calculations were performed in Excel.

An important rule to apply before simulating the logs cut from a tree stem is to determine if the tree-length is to be kept in one piece or cross-cut into two long-log lengths. Essentially, long-logs loaded onto the haul trucks need to be less than 19.5 m to meet transport regulations and the harvesting contractor will not cut long-logs shorter than 10 m in order to maximize load size. Therefore, if the top merchantable height of the standing tree is greater than 20.31 m (accounting for 0.3 m for stump height and butt-end trim) then two long-logs are cut out of a standing tree. If the standing tree is less than 20.3 m, then the tree-length is kept as a single long-log. Determining whether one or two long-logs are to be cut from a stem is important because the first log cut from a long-log is 3.17 m, whereas all remaining logs are 2.51 m, thus if two long-logs are cut from a tree, then there will be two 3.17 m logs. An example of these two different cutting patterns can be seen in Figure 12.

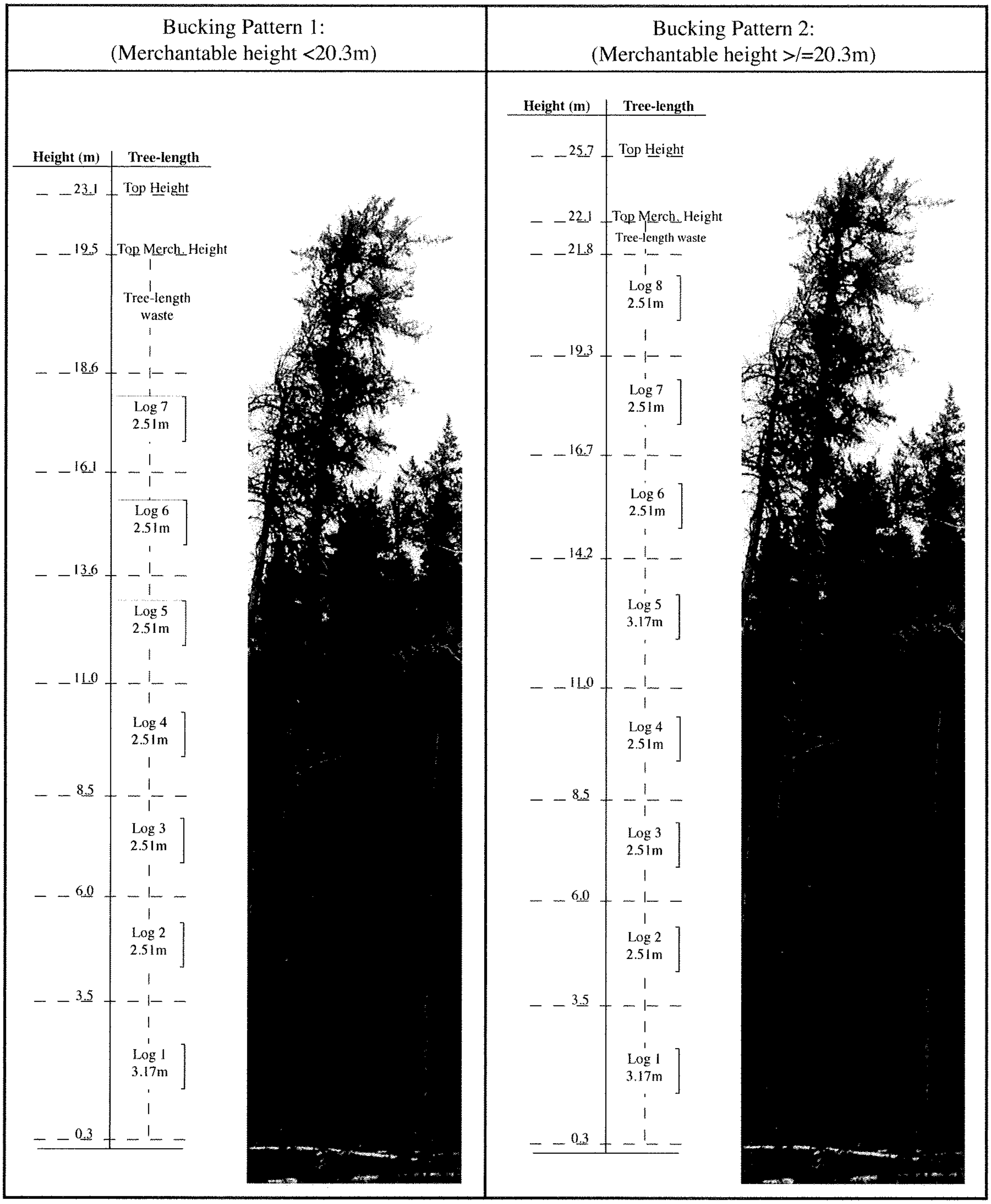


Figure 12. Tree bucking pattern 1 and 2 applied to all merchantable tree-length stems that have a minimum length of 10m and top diameter of 10cm.

3.4.2 Experimental Design

Due to the large number of predictor variables explaining lumber value and volume recovery, there is a problem of multicollinearity in conducting multiple regression analysis of log variables scanned at the sawmill. Therefore, principal component analysis (PCA) was used to isolate principal components from the original log variables. The end goal is to build regression models for predicting GLVR in the sawmill with the principal components extracted (detailed regression models are shown in Chapter 4).

One of the challenges in this study is data correlation between different data collection stages. For example, 1:1 correlation of tree-level data to the GLVR extracted from the logs sawn in the sawmill is near impossible without individual tags or signatures following each and every stem measured from forest through to lumber creation. Therefore, PCA and regression analysis were performed at the two stages: log-level (using Stage 3 data collected at sawmill) and tree-level (using Stage 1 data collected in timber cruise). PCA was performed at the log-level, and the regression models were created to predict the GLVR. Then using the bucking process described in section 3.4.1, the log-level regression model was used to calculate individual GLVR for each log extracted from the tree stem. The value recoveries for each log cut from the same stem were added together to create a GLVR estimate for the entire tree.

3.4.3 PCA at Log-level and Tree-level

The study was designed to test the hypothesis if log and tree variables measured at the sawmill and forest respectively, can be isolated into two or three principal components to be used in future regression analysis. PCA is a linear dimensionality reduction technique, which identifies orthogonal directions of maximum variance in the original data, and projects the data into a lower-dimensionality space formed of a sub-set of the highest-variance PCs (Chiorescu and Gronlund 2004; Field 2005). The PCs are derived using linear algebra. The solution is based on an important property of eigenvector decomposition. The data set is X, an m x n matrix, where m represents each external log or tree variable and n represents each log or tree observation.

Mathematically, the PC extraction goal is to find some orthonormal matrix P where Y=PX such that $S_Y = \frac{1}{n-1} YY^T$ is transformed into a diagonal matrix.

The rows of P are the principal components of matrix X (Shlens 2003). Put succinctly, PCA produces PCs from the dataset containing collinear variables. The PCs aggregate the collinear variables by representing the majority of variance found between the groups of collinear variables. The PC's essentially plot a linear regression line that best fits a group of collinear variables and assigns a regression value for each log or tree.

The analysis of variance test was used to determine the significance of each variable prior to conducting the PCA. Three preliminary statistical tests (Determinant, Kaiser-Meyer-Olkin (KMO), and Bartlett's test of sphericity) were performed to check the assumptions of singularity and non-correlation in the data. The Determinant tests extreme multicollinearity in two or more variables, KMO is a measure of sampling adequacy, and Bartlett's test of sphericity is used to test the null hypothesis of singularity. Following recommendations by Hair *et al.* (1995), factor loadings higher than 0.3 are regarded as significant. Loadings slightly less than 0.3 may, however, in some cases indicate relationships that can be discussed and analyzed. Generally, where the component plateau is the intuitive cut-off point, and when using Kaiser's criterion (Kaiser 1974) where the sample size is >250 and average communality is >0.6,all factors with eigen values above 1 should be retained for PCA.

Factor rotation is used to improve the interpretability of factors. Rotation maximizes the loading of each variable on one of the extracted factors, while minimizing the loading on all other factors (Fields 2005). There are two groups of rotation: i) orthogonal that is used for factors assumed to be unrelated, and ii) oblique that is used for factors assumed to be related. Orthogonal rotations include: varimax, quartimax, and equamax. For the PCA's in this thesis, varimax rotation was used for the log and tree-level analyses in order to maximize the dispersion of variable loadings within the principal components.

3.4.4 Multiple Regression Analysis at Log-level and Tree-level

The second stage of experimental design was to take the PC regression scores from the PC's extracted (e.g. PC1:Log Size = S and PC2: Log Form = F) from the PCA analysis and explore various linear and nonlinear multiple regression models to predict the GLVR.

The purpose of multiple regression analysis is to understand the relationship between two or more predictor variables and a dependent or criterion variables (Pearson 1908). Partial Least Square (PLS) is a regression technique commonly used in the lumber recovery models (Liu *et al.* 1989; Roos *et al.* 2000; Zhang and Tong 2005; Liu *et al.* 2007a; Liu *et al.* 2007b). PLS is a linear regression method that generalizes and combines features from PCA and multiple regression analysis (Tobias 1995; Antii 1999; Reeves and Delwiche 2003; Abdi 2003).

The lumber value and volume recovery has been determined using log size and log geometry (Roos *et al.* 2000; Via *et al.* 2003; Liu *et al.* 2007). The log-level multiple regression analysis was conducted using log size (small diameter, large diameter, volume, and length) as Principal Component 1 (PC1=S) and log form (taper, length, and sweep) as Principal Component 2 (PC2=F). The green lumber value recovery at log-level as a function of S (log size) and F (log form) is represented by equation (1) as:

$$L_j = f(S_{PC}, F_{PC}) \qquad [1]$$

Where,
L_j represents rough-green lumber value recovery (CAN$) from a log j,
S_{PC} denotes the Principal Component 1(log Size) regression score, and
F_{PC} is principal component 2 (log Form) regression score.

For the tree-level regression analysis, we use the results from the PCA analysis, where PC1(S) represents the tree size parameters (diameter profile, top merchantable height, and taper), PC2(F) the tree form parameters (taper and sweep), and PC3(B) the crown characteristics parameters (live crown ratio). The green lumber value recovery at tree level as a function of PC1 (tree size), PC2 (tree form), and PC3 (tree branchiness) is represented by equation (2) as:

$$T_j = f(S_{PC}, F_{PC}, B_{PC}) \qquad [2]$$

Where,
T_j represents rough-green lumber value recovery (CAN$) from a tree j,
S_{PC} denotes the Principal Component 1 (tree Size) regression score,
F_{PC} is principal component 2 (tree Form) regression score, and
B_{PC} is principal component 3 (tree Branchiness) regression score.

Finally, we examined three types of model forms: linear function, exponential function and power function, using least square regression method for estimating and validating the reliability of the model (Gujarati 1995).

Models were evaluated based on the coefficient of determination (R^2), the mean squared error (MSE), and the root mean squared error (RMSE) of predictions using the equations (3), (4) and (5), respectively:

$$R^2 = 1 - \frac{(P_j - \hat{P}_j)^2}{(P_j - \bar{P})^2} \qquad [3]$$

$$MSE = \frac{\sum_{j=1}^{m} |P_j - \hat{P}_j|^2}{n} \qquad [4]$$

$$RMSE = \sqrt{\frac{\sum_{j=1}^{m} |P_j - \hat{P}_j|^2}{n}} \qquad [5]$$

Where,
P_j and $\hat{P}_j$ are observed and predicted rough-green lumber value recovery of logs j,
$\bar{P}$ is the mean of the observed rough-green lumber value recovery for the n logs,
j= 1,2,…,n, and
n is the number of logs (n = 16565 in this study).

The R^2 is the square of the sample correlation coefficient between the outcomes and their predicted values (Fields 2005). R^2 is helpful in understanding the goodness of fit of the model. The MSE and RMSE are helpful to explain the quality of the model by proving an indicator of typical misfits of a model or the standard error of the estimate (Fields 2005). We use both the R^2 and MSE to assess a model because although R^2 tells us the goodness of fit of the model, MSE is used to explain the width of the confidence intervals for prediction, or the amount of error between the observed and predicted values. RMSE is useful when discussing the real error in a model as RMSE is measured in the same units as the data, rather than in squared units, and is representative of the size of average error (Fields 2005).

In deciding which models to use, an exploratory method was applied. The regression analysis started with the running of a stepwise regression analysis to see if combining different PCs as predictor variables would improve the model accuracy. Two types of nonlinear models (power and exponential functions) were also formulated and compared using the model criteria: R^2, MSE and RMSE.

3.4.5 Block-level Productivity Comparison

A block-level productivity comparison was created once the PCA and regression analyses were complete for both log and tree-level data. To generate a block-level productivity comparison the bucking simulation and tree-length profiles were combined to calculate an estimated product recovery for each stem in the timber cruise. The timber cruise information was used to extrapolate the GLVR of different stems (Model 7_T) to a per hectare level. The GLVR for an entire hectare was calculated. Different lumber recovery factors (LRFs) along the supply chain were compared by taking the final GLVR and board volume scanned at the sawmill trimmers and dividing these recovery values by total jack pine volumes recorded in the timber cruise, weigh scale, and at the log line scanners.

4. RESULTS AND DISCUSSION

4.1 LOG-LEVEL PCA

4.1.1 Preliminary Analysis of Singularity and Non-correlation

The descriptive statistics of log level variables along with their preliminary correlation information with GLVR and significance are shown in Table 10. It was seen that diameter, length and volume were strongly correlated with GLVR and have a significant effect on GLVR. Sweep and taper have little or no correlation with GLVR and only taper has a significant effect on GLVR. Similar results were also found by Steele (1984), and Moberg and Nordmark (2006) who found log quality and knot characteristics are also important. The sawmill studied did not scan for knots in the GLVR process and thus are not included in the analysis. This is a definite shortcoming of the log-level models, as we know from many researchers (Kellogg and Warren 1984; Steele 1984; Zhang and Gingras 1999; Zhang *et al.* 2001; Beauregard *et al.* 2002; Wilhelmsson and Moberg 2004; Moberg and Nordmark 2006;) that knot size, angle and distribution play a significant role in GLVR in jack pine and other conifers.

Table 10. Descriptive statistics used for establishing log-level PCA analysis (n = 16565).

Descriptive	Small End Diameter(cm)	Large End Diameter(cm)	Length(cm)	Volume(m^3)	Sweep(cm)	Taper(cm)	GLVR($)
Minimum	8.13	10.41	152.40	0.02	0.30	0.00	0.00
Maximum	40.39	58.67	335.28	0.43	13.18	28.80	53.20
Mean	19.38	21.49	266.43	0.09	2.04	2.19	9.73
SD	5.15	5.85	27.07	0.06	1.11	1.57	7.10
Linear R^2 predicting GLVR($)	0.87	0.91	0.91	0.97	<0.001	0.26	-
Anova Sig. Test	<0.001	<0.001	<0.001	<0.001	0.32	<0.001	-

The correlation matrix describes the level of collinearity between variables and tests whether or not two variables are significantly similar. The results show that there is high collinearity between the selected diameter (diameter measured 2 feet from small-end of log), small-end diameter and large-end log diameter measurements (Table 11). The final statistic in gauging whether the data has extreme multicollinearity, we use the determinant of the correlation matrix. Since the value of the determinant is less than 1.0E-5, one or more variables need to be removed (Hair *et al.* 1995; Field 2005). We removed the variable Selected Diameter, as shown in the correlation matrix for culled log variables in Table 12. The new determinant value of 1.28E-05, is greater than the threshold value of 1.0E-5, proving that singularity is not a problem in the revised dataset.

Table 11. PCA correlation matrix for all logs variables reporting correlation coefficients, tests of significant similarity, determinant, KMO, and Bartlett's.

		Selected Diameter	Small End Diameter	Large End Diameter	Length	Volume	Sweep	Taper
Correlation	Selected Diameter	1.000	0.997	0.972	0.350	0.943	0.059	0.365
	Small End Diameter	0.997	1.000	0.970	0.347	0.940	0.055	0.348
	Large End Diameter	0.972	0.970	1.000	0.464	0.973	0.103	0.564
	Length	0.350	0.347	0.464	1.000	0.549	0.205	0.611
	Volume	0.943	0.940	0.973	0.549	1.000	0.103	0.563
	Sweep	0.059	0.055	0.103	0.205	0.103	1.000	0.212
	Taper	0.365	0.348	0.564	0.611	0.563	0.212	1.000
Sig. (1-tailed)	Selected Diameter	-	0.000	0.000	0.000	0.000	0.000	0.000
	Small End Diameter	0.000	-	0.000	0.000	0.000	0.000	0.000
	Large End Diameter	0.000	0.000	-	0.000	0.000	0.000	0.000
	Length	0.000	0.000	0.000	-	0.000	0.000	0.000
	Volume	0.000	0.000	0.000	0.000	-	0.000	0.000
	Sweep	0.000	0.000	0.000	0.000	0.000	-	0.000
	Taper	0.000	0.000	0.000	0.000	0.000	0.000	-

Determinant = 3.27E-08
KMO = .708
Bartlett's = p <0.001

Table 12. PCA correlation matrix for culled logs variables reporting correlation coefficients, tests of significant similarity, determinant, KMO, and Bartlett's.

		Small End Diameter	Large End Diameter	Length	Volume	Sweep	Taper
Correlation	Small End Diameter	1.000	0.970	0.347	0.940	0.055	0.348
	Large End Diameter	0.970	1.000	0.464	0.973	0.103	0.564
	Length	0.347	0.464	1.000	0.549	0.205	0.611
	Volume	0.940	0.973	0.549	1.000	0.103	0.563
	Sweep	0.055	0.103	0.205	0.103	1.000	0.212
	Taper	0.348	0.564	0.611	0.563	0.212	1.000
Sig. (1-tailed)	Small End Diameter	-	0.000	0.000	0.000	0.000	0.000
	Large End Diameter	0.000	-	0.000	0.000	0.000	0.000
	Length	0.000	0.000	-	0.000	0.000	0.000
	Volume	0.000	0.000	0.000	-	0.000	0.000
	Sweep	0.000	0.000	0.000	0.000	-	0.000
	Taper	0.000	0.000	0.000	0.000	0.000	-

Determinant = 1.28E-05
KMO = .588
Bartlett's = p <0.001

The final test in PCA to ensure that the PCA will yield acceptable results is the KMO test that measures sampling adequacy (Field 2005). In the PCA with Selected Diameter removed, KMO is 0.588 (Table 13), which shows an acceptable pattern of correlation between the variables and should yield distinct factors (Kaiser 1974; Field 2005). The Bartlett's test of sphericity result is shown in Table 13. The p-value is <0.001, therefore, we reject the null hypothesis that the variables are not correlated proving that there are some correlations between the variables that can be used to perform the PCA.

Table 13. KMO measure of sampling adequacy and Bartlett's test of sphericity.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.588
Bartlett's Test of Sphericity	Approx. Chi-Square	196787.459
	df	15
	Sig.	0.000

4.1.2 Factor Extraction

The total variance explained by components is shown in Table 14. Two components were extracted with eigen values greater than 1; meaning they explain more variance in data than any single variable. We see that the un-rotated solution shows the first and second components explain 61.8 % and 17.8 % of the variance, respectively. Reducing a large list of log variables into components to be linked through multiple regression to GLVR was not found in the body of literature relating to lumber recovery from logs studies. This is probably due to the common use of sawing simulators to model lumber recovery (Grondin and Drouin 1998; Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a and 2007b; Kantola *et al.* 2008 and 2009) at the sawmill thus reducing the need to model the GLVR from logs back to standing trees using PCA and regression. However, PCA has been used in lumber processing for tracking boards by compressing a 3D image analysis into a 2D space (Bharati and MacGregor 2003). Flodin *et al.* (2008) used PCA to reduce several log factors (length, top taper, butt taper, sweep height, sweep radius, sweep position) into PC to track logs through the supply chain. Neither of these researchers used PCA to generate regression scores for use in multiple regression.

In Table 14 the total variance explained in the rotated solution remains the same as in the un-rotated solution. The difference between the un-rotated and rotated solutions is the % of variance explained by each component has changed trending towards a more equalized proportion of the extracted variance. Component 1 and 2 now explain 56.2 % and 23.4 %, respectively. The reason for the change in variance explained by each factor is shown in Table 16, where Taper, Sweep and Length show a higher correlation with component 2 in the rotated solution than in the un-rotated solution.

Table 14. Total variance of the culled log-level dataset explained by components with extraction sum of squares loadings and rotated sum of squares loadings.

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.631	60.52	60.52	3.631	60.52	60.52	3.192	53.19	53.19
2	1.152	19.21	79.73	1.152	19.21	79.73	1.592	26.54	79.73
3	.793	13.22	92.95						
4	.395	6.59	99.54						
5	.027	.46	100.00						
6	.000	.00	100.00						

The scree plot in Figure13 also represents the contribution of each component in explaining the variance in the data. Figure 13 shows that component 1 explains the majority of the variance in the data and the subsequent components tend to plateau. Generally, where the components plateau is the intuitive cut-off point (Kaiser 1974; Field 2005), however, all factors with eigen values above 1 should be retained for PCA when using Kaiser's criterion where the sample size is >250 and average communality is >0.6. Therefore, we retained components 1 and 2 for further analysis.

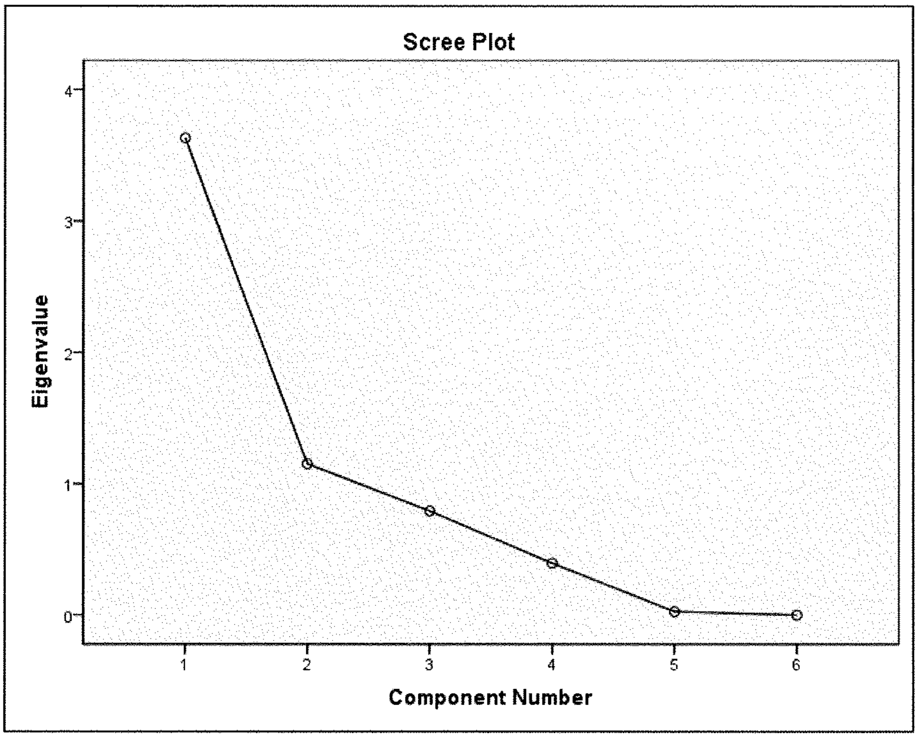


Figure 13. Scree plot showing eigen value trend by component number.

The proportion of variance extracted from each log variable is shown in Table 15. The proportion of variance extracted is good, with the highest value being 0.977 and the lowest value being 0.602 for Large End Diameter and Sweep, respectively. An average extraction value of 0.797 is well above the 0.6 level of acceptability. Therefore, the PCA has a good level of extraction (Kaiser 1974; Field 2005). Although Flodin *et al.* (2008) used PCA to reduce a larger number of log variables into components, they did not report the extraction values and therefore a comparison cannot be made. Other researchers using PCA (Wold *et al.* 1987; Eriksson *et al.* 2001; Chiorescu 2003) have found extractions greater than 0.7 to be acceptable for PCA.

Table 15. Communalities in the variables explaining the proportion of data variance within each variable explained by the principal component factors.

	Communalities	
	Initial	Extraction
Small End Diameter	1	0.929
Large End Diameter	1	0.977
Length	1	0.638
Volume	1	0.975
Sweep	1	0.602
Taper	1	0.663
Average	1	0.797

Extraction Method: Principal Component Analysis.

4.1.3 Factor Rotation and Interpretation

The Rotated Component Matrix (b) in Table 16 presents the clearer grouping of variables by component. Although Flodin *et al.* (2008) used PCA in a log-level analysis, the component plots were not shown in the report. There are no other PCA studies found in the literature that provide a useful log-level comparison of the component plots. Table 16 shows that diameter and volume measurements are strongly loaded on factor 1. Sweep, a type of form, is strongly loaded on factor 2. Length and taper have significant loadings on both factor 1 and 2 and, therefore, are expressed as a combination of factor 1 (log size) and factor 2 (log form). These groupings of variables are shown in an un-rotated loading plot in Figure 14. A second loading plot is shown in Figure 15, where the plot is projected using a varimax rotation. The varimax rotation shows a clearer grouping of variables by PCs. Varimax rotation was also found to show the clearest grouping of variables by several researchers (Eriksson *et al.* 2001; Chiorescu 2003: Field 2005; Flodin *et al.* 2008).

Table 16. Variable loadings by each principal component for the component matrix (a) and rotated component matrix (b) with values <0.3 (or >-0.3 if negative) excluded.

	Component Matrix(a)		Rotated Component Matrix(b)	
Variable	1	2	1	2
Large End Diameter	0.962	-	0.968	-
Small End Diameter	0.886	-0.379	0.964	-
Volume	0.97	-	0.957	-
Sweep	-	0.749	-	0.765
Length	0.668	0.438	0.421	0.679
Taper	0.702	0.412	0.463	0.669

(a) Extraction Method: Principal Component Analysis.

(b) Rotation Method: Varimax with Kaiser Normalization.

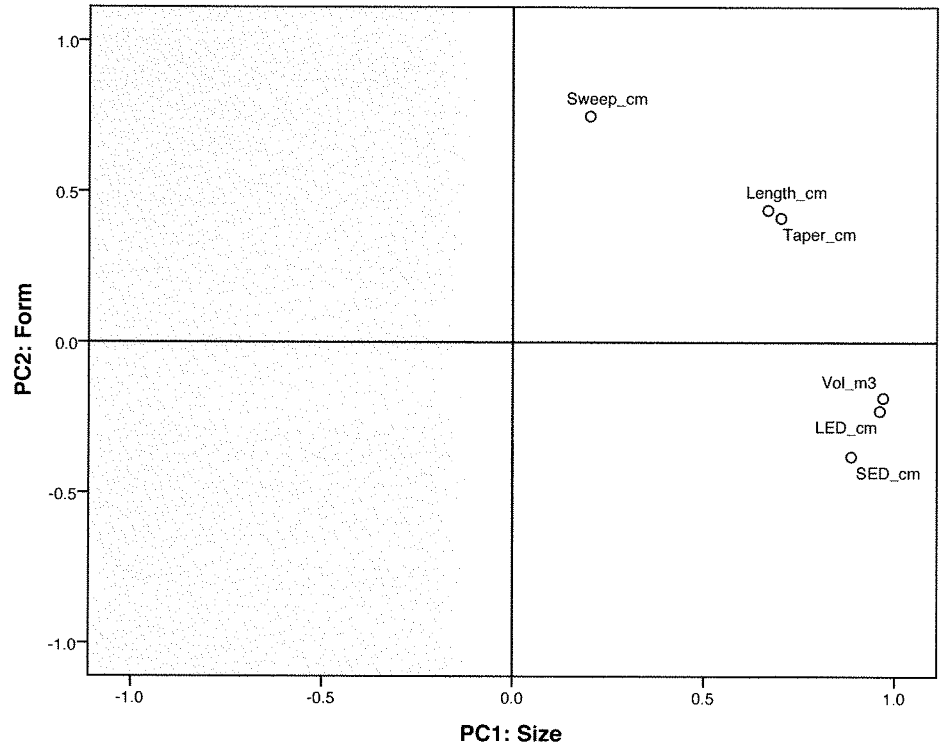


Figure 14. Log variables scores plotted by principal components (unrotated).

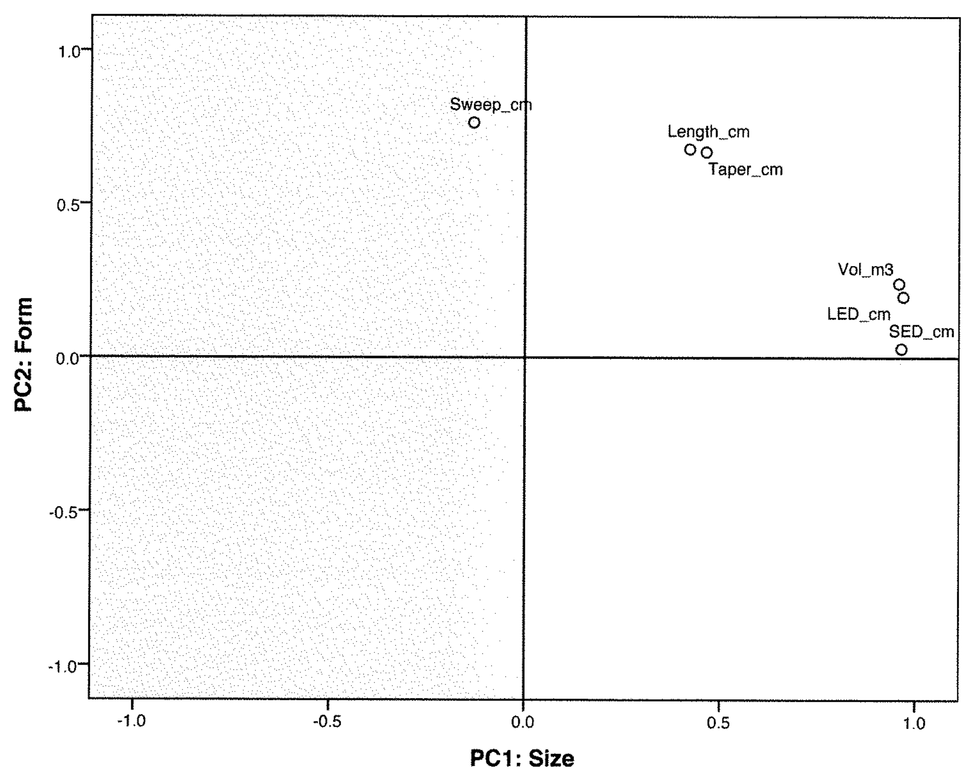


Figure 15. Log variables plotted against component scores in varimax-rotated space.

We can see from Figures 14 and 15 that once rotated, the variables loaded on axis 1 and axis 2 clearly distinguish two names for the components. Component 1 can be termed as Log Size, and Component 2 as Log Form. We see that PC1explains size components (log volume, small diameter and large diameter), and PC2 explains form variables (sweep). Taper and length variables are explained by an interaction of size and form components. The literature does not appear to have a comparable study to this thesis, where log-level PCA was used to reverse-forecast GLVR in individual stems. PCA was used in log tracking and image analysis but none of the components were reported therefore there is no point of comparison to other log-level studies.

4.2 LOG-LEVEL REGRESSION

4.2.1 Model Development

Table 17 presents the summary statistics for the 16,565 sample logs run through the sawmill. Taking the results produced from the log-level PCA analysis in section 4.1, a regression analysis was conducted to determine if the principal components extracted from log level variables can be used to predict the GLVR from sawlogs. Several models were created, tested and compared. Figure 16 shows the GLVR ($) by log size ($m^3$). This is the data trend we are striving to model using PC1:Size and PC2:Form. The modeling technique of using PC regression scores as independent variables in multiple regression for modeling GLVR does not appear in the literature. This may be due to the easier comprehension of using measured variables as the predictor variables. However, previous researchers have found that using PC regression scores as either tracking signatures or independent variables explain a higher level of variance than any one measured variable (Bharati and MacGregor 2003; Flodin *et al.* 2008). The fact that PCs can represent more of the data variance than any one measured variable (Field 2005), and thus improve model performance with fewer variables, was one of the primary reasons for this modeling approach. Additionally, most lumber recovery studies use scans of the log to run optimization routines in lumber recovery (Steele 1984; Middleton *et al.* 1989; Liu *et al.* 1989; Shi *et al.* 1990; Wagner and Taylor 1993; Roos *et al.* 2000; Zhang and Tong 2005; Liu *et al.* 2007a; Liu *et al.* 2007b).

Table 17. Summary statistics of the data set used for establishing log-level regression models (n=16565).

	PC1: Size	PC2: Form	GLVR ($/log)
Minimum	-2.71	-2.98	0.00
Maximum	5.03	9.55	53.20
Mean	0.00	0.00	9.73
SD	1	1	7.10

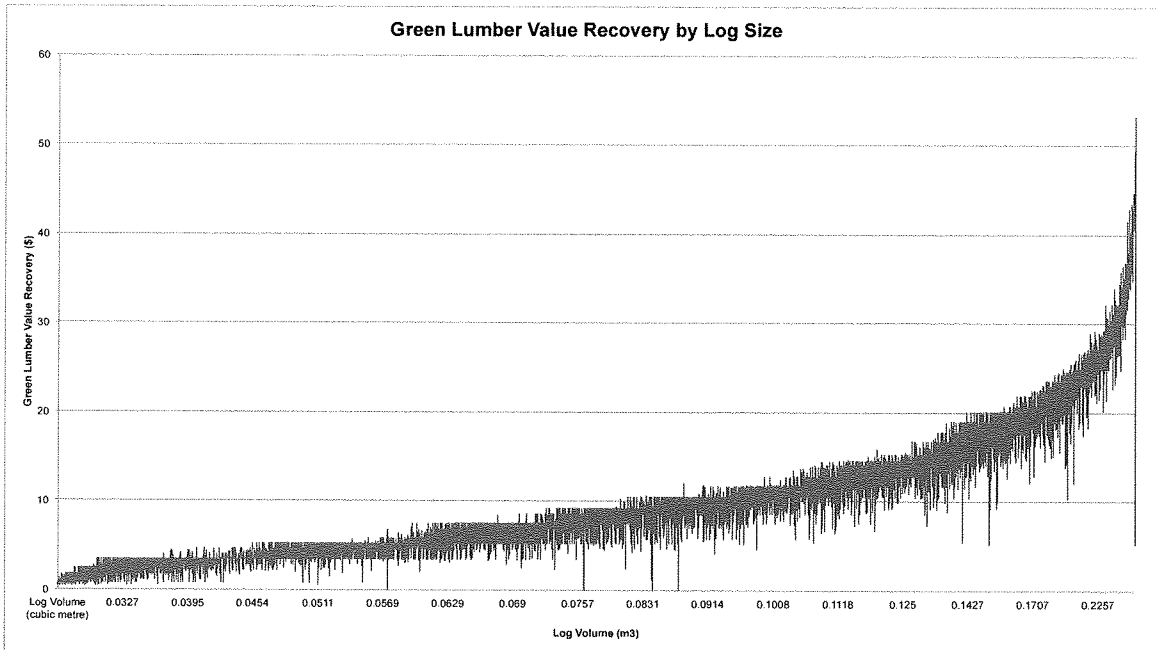


Figure 16. Measured GLVR ($) by average log size ($m^3$/log).

Table 18 shows the summarized results from the PCA analysis and how each PC is linearly correlated with the GLVR. The linear correlation with GLVR shows a R^2 value of 0.936 with PC1 (Size as a predictor variable) and a R^2 of 0.021 with PC2 (Form as a predictor variable). The use of log-level regression correlations of PCs to predict GLVR was not found in the review of the literature. The tree-level modeling in sections 4.3 and 4.4 have results that can be compared to the literature. Log-level modeling in the literature appears to focus on log tracking (Flodien *et al.* 2008) and log image analysis

(Bharati and MacGregor 2003). From this preliminary look at correlations with the GLVR, it is apparent that GLVR has the strongest correlation to log size and to a much lesser extent, log form. Furthermore, the log diameters (both small and large) and log volume are the most important variables in determining a log's GLVR. Figure 17 illustrate the strong fit of PC1 with the GLVR. Figure 18 shows a poor linear fit between PC2 and the GLVR, however, the ANOVA test of significance in Table 18, reveals that the PC2:Form has a significant effect in predicting the GLVR, thus it needs to be included in further analysis. The observation that size and form of logs are significant factors in determining log GLVR is supported by many studies (Kellogg and Warren 1984; Steele 1984; Zhang and Gingras 1999; Beauregard *et al.* 2002; Wilhelmsson and Moberg 2004; Moberg and Nordmark 2006; Liu *et al.* 2007a, Liu *et al.* 2007b). These studies also included knot size and frequency as significant, but the sawline scanners used in the study did not include knots and are therefore omitted from the analysis.

Table 18. Summarized results of log-level principal component analysis used in regression modeling.

		Principal Componenets Extracted	
		PC1: Size	PC2: Form
Eigenvalue		3.631	1.152
Proportion of total variance explained (%)		60.524	19.206
Loadings (>0.3)	Large End Diameter	0.968	-
	Small End Diameter	0.964	-
	Volume	0.957	-
	Sweep	-	0.765
	Length	0.421	0.679
	Taper	0.463	0.669
Linear R^2 predicting GLVR($)		0.936	0.021
Anova Sig. Test		<0.001	<0.001

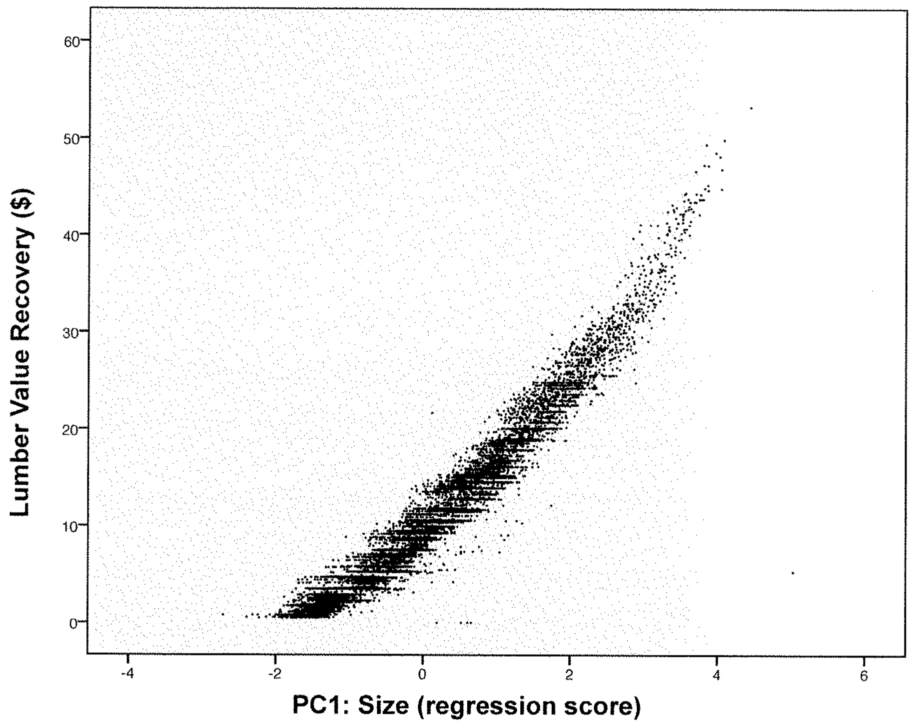


Figure 17. Plots of principal component 1 regression scores against GLVR ($/log).

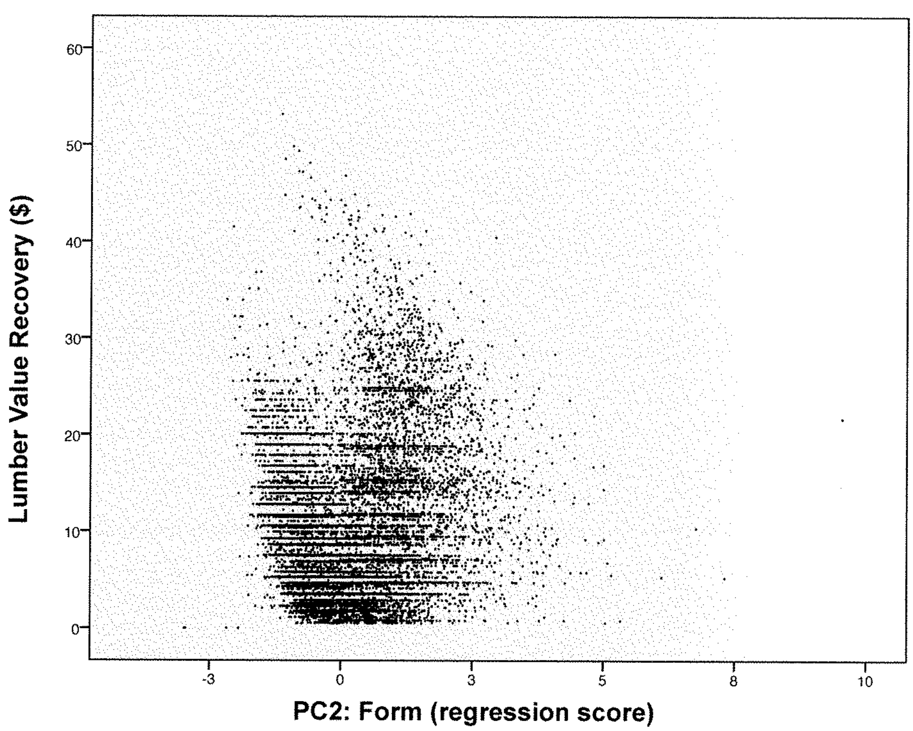


Figure 18. Plots of principal component 2 regression scores against GLVR ($/log).

Three types of models were tested in the analysis. These types were linear functions, non-linear power functions, and non-linear exponential functions. The use of PC regression scores in modeling GLVR was not found in the review of the literature, however, similar studies modeling tree-level GLVR were performed by several researchers (Zhang *et al.* 2001; Zhang and Tong 2005; Zhang *et al.* 2006; Liu et al. 2007a; Liu et al. 2007b) who used linear and non-linear functions. How these studies compared with the thesis are discussed in more detail in section 4.4. A total of seven individual models were run and the model forms are shown in Table 19 with the parameter estimates and criteria shown in Table 20. Residual plots are shown in Figures 19 and 20.

Table 19. Model forms for estimating green lumber value recovery using log-level principal components size (S) and form (F): L is GVLR in $/log, and a_0, a_1, a_2 and a_3 are constant coefficients).

Model number	Model form
1_L	$L = a_0 + a_1 * S$
2_L	$L = a_0 + a_1 * S + a_2 * F$
3_L	$L = a_0 + a_1 * S^2 + a_2 * F^2$
4_L	$L = a_0 + a_1 * S^3 + a_2 * F^3$
5_L	$L = \exp(a_0 + a_1 * S + a_2 * F)$
6_L	$L = a_0 + \exp(a_1 * S + a_2 * F)$
7_L	$L = a_0 + a_1 * F + a_2 * \exp(a_3 * S)$

4.2.2 Model Comparison and Evaluation

The results of the model parameter estimates and criteria are shown in Table 20. The first nonlinear models run were power functions. The residual plots in Figure 19 and criteria in

Table 20 show that both power functions performed poorer than the linear models. The third model type employed was an exponential model. Models 5_L to 7_L show different exponential models. Model 5_L shows a better higher-order model than the power models, however, not as good a fit as the linear models. Models 6_L and 7_L are modifications of the exponential model, where terms are moved from the exponential function to linear components of the model preceding the exponential function. Model 7_L, which uses the exponential form of PC1:Size, adjusted by the linear expression of PC2:Form, produced the best fit of the data. The purpose of attempting different models with both PC1:Size and PC2:Form was to assess the interaction and combination of the two significant factors to find the best modeling of the GLVR. The exponential models fit the data the strongest because we see that with a linear increase of volume, there is an exponential increase in value (Figure 16). Tree-level models have been created by other researchers and have a similar exponential trend in value but the strongest models were polynomial functions (Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a). Further discussion of these tree-level models is found in section 4.4.

The first higher order functions tested was power functions. Model 3_L was used as a 2^{nd} order power function, which yielded the lowest R^2 value of 0.406 and the highest MSE value of 29.92. Model 4_L shows a third order power function with an improved fit compared to Model 3_L with an R^2 of 0.584 and a MSE 209.98. Neither power functions performed better than the linear functions, therefore, a third type of function needed to be explored. Liu *et al.* (2007a) also found that power function models did not perform as well as linear models.

The three types of function tested were exponential functions. Model 5_L included both PC1 and PC2 in an exponent function, yielding an R^2 of 0.903 and a MSE of 4.87.

The residual plots are shown in Figure 19 and show a lack of fit. Therefore other forms of the exponential model needed to be explored. Model 6_L is a modified version of Model 5_L to create a linear intercept coefficient before the exponential function of PC1 and PC2. Model 6_L yielded an R^2 of 0.544 and a MSE of 22.96. Figure 19 further illustrates that Model 6_L is less fitted than exponential model 5_L or both linear models.

A third exponential model was attempted which combined characteristics of the linear model 2_L and exponential features. The reasoning for attempting this model was based on the linear models (1 and 2) where PC2:Form contributed very little to the overall model R^2 but seemed to improve the MSE more, in terms of a ratio. Therefore, Model 7_L takes the linear function of PC2:Form combined with the nonlinear exponential function of PC1:Size resulting in an R^2 of 0.972 and a MSE of 1.42. Zhang and Tong 2005, Zhang *et al.* 2006, Liu *et al.* 2007a and Liu et al. 2007b also found that combining linear and non-linear functions in a model improved performance. Model 7_L shows the best model fit to predict the GLVR with the log-scan data recorded at the sawmill. Model 7_L predicted values overlapped by measured values are shown in Figure 21(ii). Figure 21(i) represents Model 7_L without PC2(log form) in the model to show the affect PC2 has on model fit. The modeling process described for the log-level data is useful because it shows a way to link external log measurements to the GLVR before processing. The utility of the selected log-level models will be further increased in sections 4.3 and 4.4 where it is used to generate base-line GLVR values for individual tree measured pre-harvest.

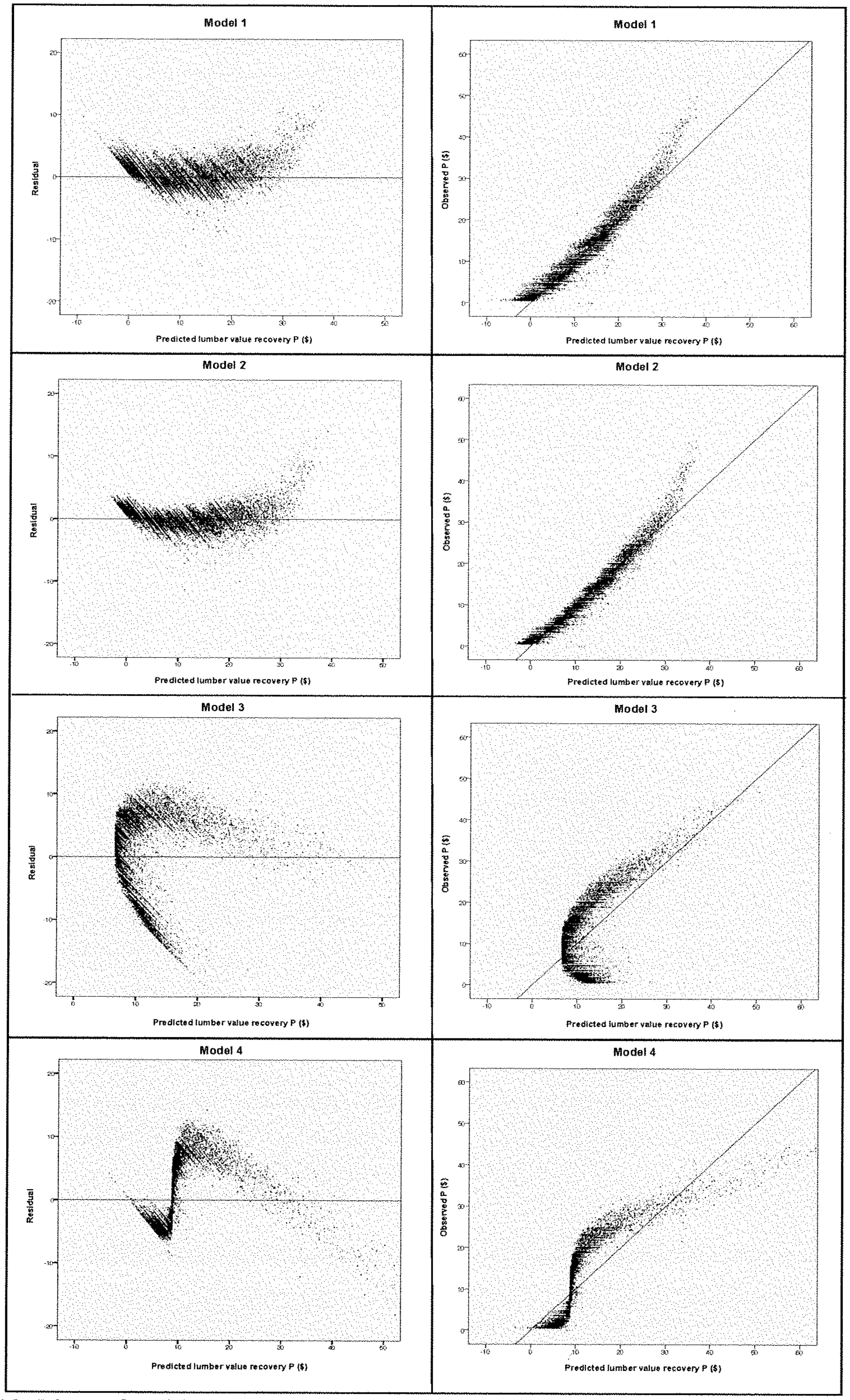


Figure 19. Plots of residuals against predicted lumber value recovery P ($/log) and observed lumber value against predicted GLVR in Models 1_L to 4_L.

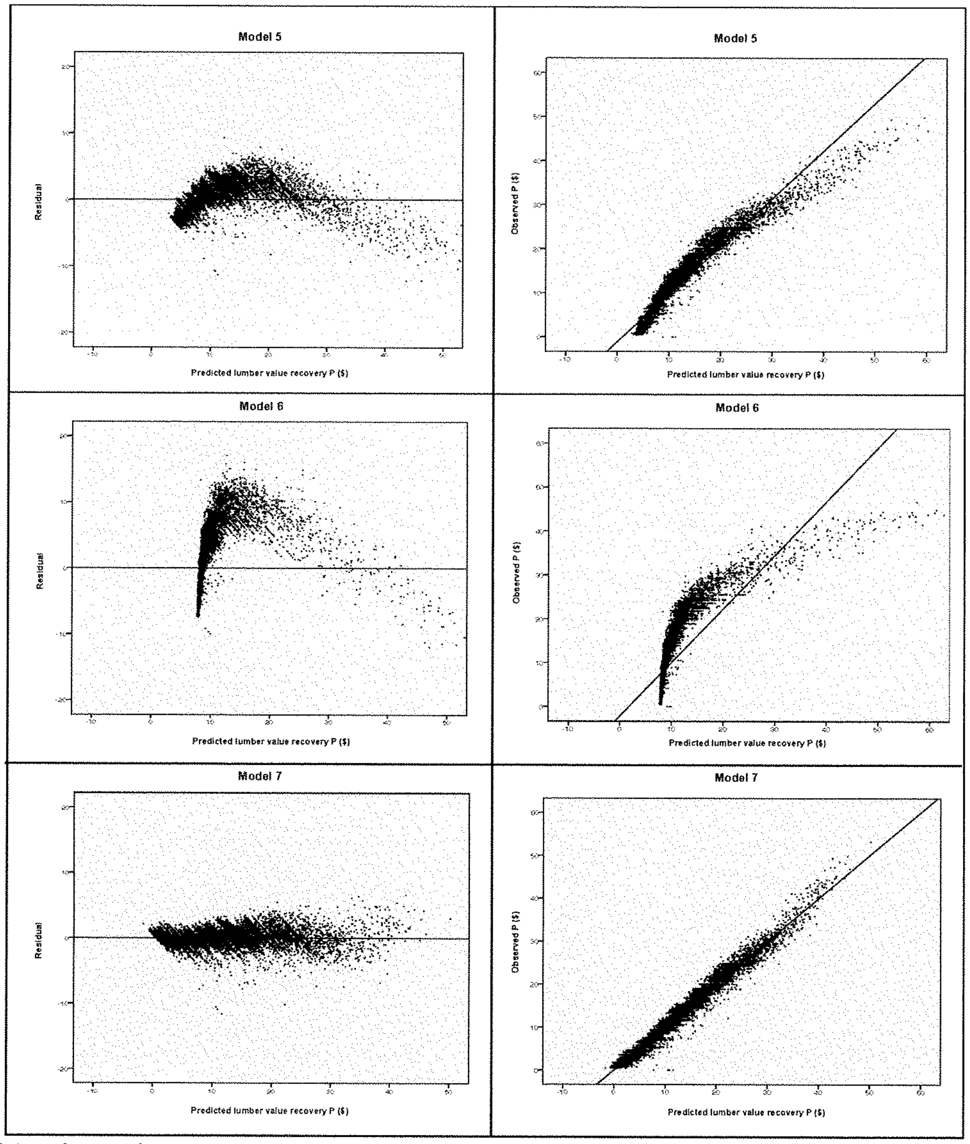


Figure 20. Plots of residuals against predicted lumber value recovery P ($/log) and observed lumber value against predicted lumber value in the sawmill (Models 5_L to 7_L).

Table 20. Parameter estimates and statistical criteria for the 7 log-level regression models.

Model	Parameters				Criteria		
number	a_0	a_1	a_2	a_3	R^2	MSE	RMSE
1_L	9.731	0.967	-	-	0.936	3.229	1.344
2_L	9.731	0.967	0.145	-	0.957	2.164	1.013
3_L	6.579	2.681	0.471	-	0.406	29.924	4.537
4_L	8.812	1.077	0.029	-	0.584	20.981	3.701
5_L	1.000	0.015	-0.204	-	0.903	4.874	1.651
6_L	7.699	1.047	-0.065	-	0.544	22.964	3.835
7_L	-27.467	0.864	36.625	0.172	0.972	1.420	0.797

Above all, the principal components PC1:Size and PC2:Form extracted (shown in Section 4.1), were successfully able to produce linear and nonlinear models to predict the GLVR in the sawmill studied. The exponential function seen in Model 7_L predicted the GLVR most accurately. However, the linear model (Model 2_L) also performed very well. PC1:Log Size was the most important factor predicting the GLVR and PC2:Log Form was significant but contributed less to predict the GLVR. Zhang and Tong (2005) and Liu *et al.* (2007a and 2007b) found that stem size as a function of diameter and height as more significant than taper in predicting lumber value recovery, thus supporting the similar trend of log size being more important than log form presented in the model above.

$L_j = 27.467 + 0.0864{*}F_{PC2} + 36.626{*}\exp(0.172{*}S_{PC1})$ [Model 7_L]

Where,
L_j represents GLVR (CAN$) from a log j,
S_{PC1} denotes the regression score for PC1(Log Size) extracted from the log-level variables (small end diameter, large end diameter, volume, length, taper and sweep) in the PCA for each log.
F_{PC2} denotes the regression score for PC2(Log Form) extracted from the log-level variables (small end diameter, large end diameter, volume, length, taper and sweep) PCA for each log.

The lumber value estimation procedure for logs utilizes the PCA to extract PC regression scores from 6 external log variables (small end diameter, large end diameter, volume, length, taper and sweep). These regression scores of PC1:Size and PC2:Form are used as the dependent variables in the nonlinear exponential multiple regression formula (Model 7_L) to estimate log-level GLVR (Figure 21). This formula was used to calculate individual lumber recovery values for the logs extracted through simulation (procedure described in section 3.4.1) from pre-harvest inventory to assign an overall GLVR for each stem. This assigned GLVR for each stem was used in the tree-level PCA and regression analyses.

Researchers have done a similar analysis using measured variables for tree-level data (Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a; Liu *et al.* 2007b). The modeling process described for the log-level data is useful because it shows a way to link external log measurements to its GLVR before processing.

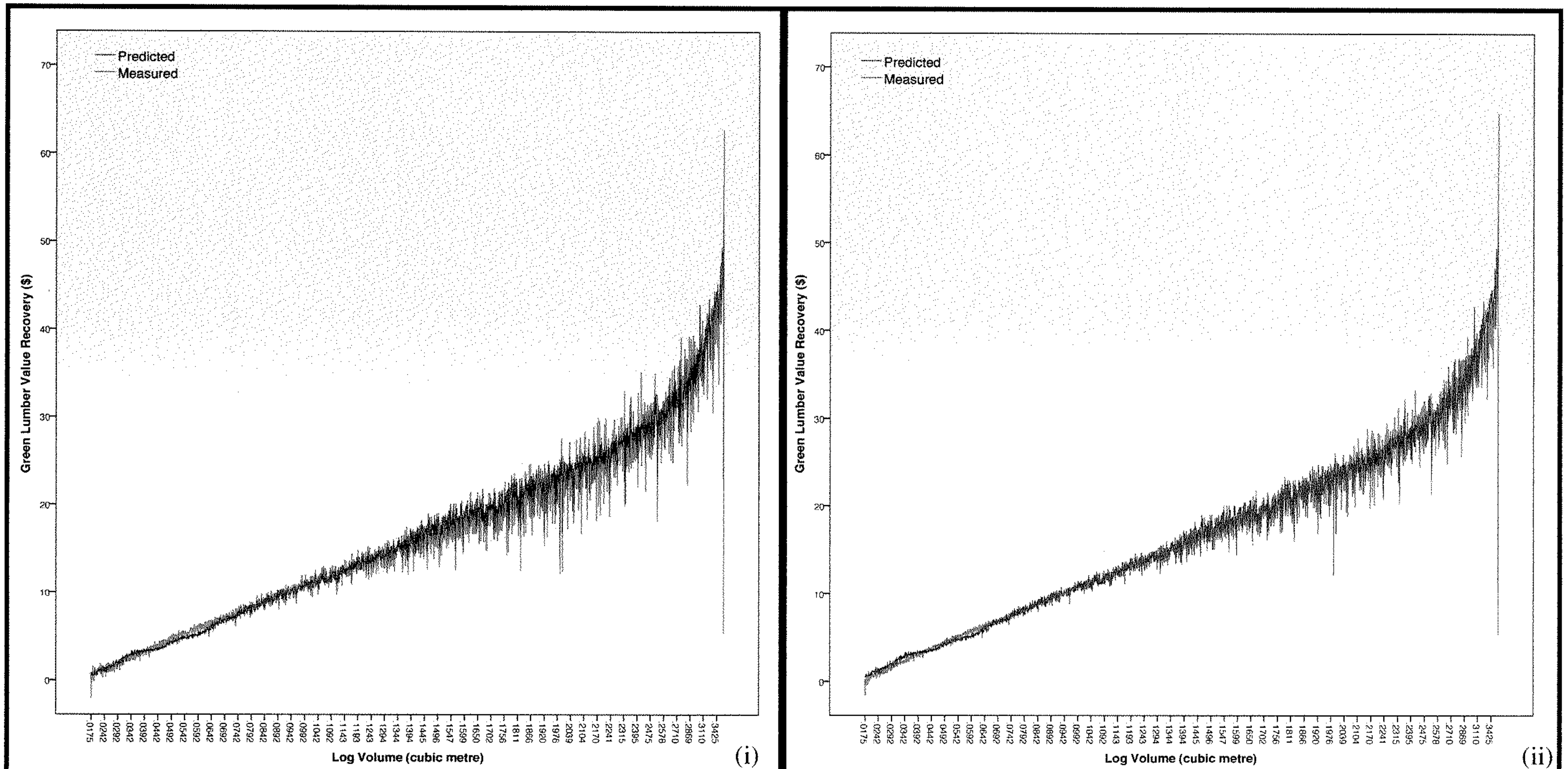


Figure 21. Observed GLVR by log size over-layed with model prediction using PC1(i) and PC1 and PC2(ii) in exponential regression models.

4.3 TREE-LEVEL PCA

4.3.1 Preliminary Analysis of Singularity and Non-correlation

The descriptive statistics of tree level variables along with their preliminary correlation information with GLVR and significance are shown in Table 21. It was seen that diameter profile and merchantable height have a strong linear correlation with GLVR and have a significant effect on GLVR. Taper has a weak linear correlation with GLVR and has a significant effect on GLVR. Live Crown Ratio and Sweep have a very weak linear correlation with GLVR and do not have a significant effect on GLVR.

Table 21. Descriptive statistics used for establishing tree-level PCA analysis (n = 101).

Descriptive	d_0.3	d_1.3	d_5	d_10	d_15	d_20	ht_merch	Live Crown Ratio	Sweep_sum	Taper	GLVR
	(cm)	(cm)	(cm)	(cm)	(cm)	(cm)	(m)	(%)	(cm)	(cm)	$
Minimum	17.44	16.00	14.00	9.00	0.00	0.00	10.00	7.44	12.50	2.50	11.16
Maximum	52.90	45.00	40.00	32.00	28.00	22.00	26.00	42.90	100.00	37.95	164.85
Mean	33.62	28.90	25.71	21.25	16.10	4.40	19.13	23.62	38.12	12.73	66.73
SD	6.45	5.41	4.99	4.63	5.28	6.82	2.70	6.45	11.46	7.10	11.16
Linear R^2 predicting GLVR($)	0.735	0.823	0.889	0.896	0.760	0.515	0.562	<.001	0.001	0.103	-
Anova Sig. Test	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.972	0.789	0.001	-

The analysis for the tree-level data follows the same approach and work flow as the PCA for tree-level data. A complete list of the variables used in the PCA is found in Table 22. As in the log-level PCA, we need to ensure the data does not have extreme multi-collinearity using the Determinant value test. Run_1 started with all variables that yielded a determinant value of less than 1.0E-5. Therefore variables with extreme multi-collinearity needed to be removed (Kaiser 1974; Field 2005). Variables selected for removal were ones with two or more correlations greater than 0.9 and were not significantly different from those variables. Using the sensitivity approach, we removed only a few variables at a time (Field 2005). Selected Run had an acceptable Determinant value of 3.10E-5, thus proving extreme multi-collinearity was not a problem in the dataset (Kaiser 1974; Field 2005).

Table 22. Variables used in each PCA run with determinant, KMO, and Bartlett values.

Tree-level Variable	Run_1	Run_2	Run_3	Run_4	Run_5	Run_6	Selected
d_0.3(cm)	x	x	x	x	x	x	x
d_1.3(cm)	x			x	x		x
d_2.5(cm)	x					x	
d_5.0(cm)	x	x	x		x		x
d_7.5(cm)	x						
d_10.0(cm)	x	x	x	x	x	x	x
d_12.5(cm)	x	x	x				
d_15.0(cm)	x	x	x	x	x	x	x
d_17.5(cm)	x	x	x				
d_20.0(cm)	x	x	x	x	x	x	x
d_22.5(cm)	x	x					
ht_merch(m)	x			x	x	x	x
ht_top(m)	x			x	x	x	
VT(cu.m/stem)	x						
VM(cu.m/stem)	x						
kh_1dead(m)	x	x	x				
kh_1live(m)	x	x	x	x	x	x	
Taper(cm/m)	x			x			
Taper(cm)	x	x	x	x	x	x	x
TaperAve(m)	x		x				
LiveCrownRatio(%)	x						x
DeadCrownRatio(%)	x						
TotalCrownRatio(%)	x						
LogCount(#)	x	x		x		x	
VMLogs(cu.m)	x						
Sweep_sum(m)	x	x		x	x	x	x
SweepAveLogs(m)	x	x	x				
SweepCodeSum	x						
SweepCodeAve	x						
Determinant	Not positive definite	1.76E-06	3.81E-06	3.67E-06	6.85E-06	3.55E-06	3.10E-05
KMO	-	0.721	0.715	0.806	0.87	0.811	0.866
Bartlett's (p-value)	-	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Table 23 shows the correlation matrix and tests of significance difference between variables for the Selected Run introduced in Table 21. We can see that live crown ratio and sweep are not significantly different from the majority of variables. However, the determinant score of 3.1E-05 allows the analysis to proceed with confidence that extreme multi-collinearity does not exist in our dataset (Kaiser 1974; Field 2005).

Table 23. PCA correlation matrix for Selected Run tree variables reporting correlation coefficients, tests of significant similarity and determinant.

		d_0.3	d_1.3	d_5	d_10	d_15	d_20	ht_merch	Live Crown Ratio	Sweep_sum	Taper
Correlation	d_0.3	1.000	0.934	0.891	0.873	0.768	0.502	0.606	0.036	0.046	0.379
	d_1.3	0.934	1.000	0.940	0.906	0.788	0.547	0.625	0.061	0.042	0.356
	d_5	0.891	0.940	1.000	0.938	0.815	0.583	0.669	0.007	0.052	0.397
	d_10	0.873	0.906	0.938	1.000	0.871	0.622	0.721	0.056	0.091	0.406
	d_15	0.768	0.788	0.815	0.871	1.000	0.649	0.862	0.062	0.158	0.352
	d_20	0.502	0.547	0.583	0.622	0.649	1.000	0.688	-0.096	0.008	0.233
	ht_merch	0.606	0.625	0.669	0.721	0.862	0.688	1.000	0.064	0.219	0.300
	LiveCrownRatio	0.036	0.061	0.007	0.056	0.062	-0.096	0.064	1.000	-0.085	0.026
	Sweep_sum	0.046	0.042	0.052	0.091	0.158	0.008	0.219	-0.085	1.000	0.188
	Taper	0.379	0.356	0.397	0.406	0.352	0.233	0.300	0.026	0.188	1.000
Sig. (1-tailed)	d_0.3	-	0.000	0.000	0.000	0.000	0.000	0.000	0.360	0.323	0.000
	d_1.3	0.000	-	0.000	0.000	0.000	0.000	0.000	0.272	0.339	0.000
	d_5	0.000	0.000	-	0.000	0.000	0.000	0.000	0.474	0.301	0.000
	d_10	0.000	0.000	0.000	-	0.000	0.000	0.000	0.289	0.182	0.000
	d_15	0.000	0.000	0.000	0.000	-	0.000	0.000	0.268	0.057	0.000
	d_20	0.000	0.000	0.000	0.000	0.000	-	0.000	0.170	0.468	0.010
	ht_merch	0.000	0.000	0.000	0.000	0.000	0.000	-	0.264	0.014	0.001
	LiveCrownRatio	0.360	0.272	0.474	0.289	0.268	0.170	0.264	-	0.199	0.400
	Sweep_sum	0.323	0.339	0.301	0.182	0.057	0.468	0.014	0.199	-	0.030
	Taper	0.000	0.000	0.000	0.000	0.000	0.010	0.001	0.400	0.030	-

Determinant = 3.1E-05

The final test is to ensure that the PCA will yield acceptable results is the KMO test that measures sampling adequacy (Field 2005). In the PCA with 15 of the original variables removed, KMO is 0.866 (Table 24), which shows a strong pattern of correlation between variables and will yield distinct factors. Other researchers using PCA (Wold *et al.* 1987; Eriksson *et al.* 2001; Chiorescu 2003) have found extractions greater than 0.7 to be acceptable for PCA. The Bartlett's test of sphericity p-value is less than 0.001 (Table 24).

Therefore, we reject the null hypothesis that the variables are not correlated and that there are some correlations between the variables that can be used to perform the PCA (Field 2005).

Table 24. KMO measure of sampling adequacy and Bartlett's test of sphericity for tree-level variables.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.866
Bartlett's Test of Sphericity	Approx. Chi-Square	994.992
	df	45
	Sig.	0.000

4.3.2 Factor Extraction

In Table 25, the total variance is explained by three components. Therefore, three components were extracted with eigen values greater than 1. We see that the un-rotated solution shows extracted percent of total data variance of 57.5 %, 11.4 % and 10.1 % for components 1, 2, and 3, respectively. A varimax rotation was used to maximize the loading of each variable to one component while minimizing its loading on the others (Field 2005). Through the use of varimax rotation, the extracted percent of total data variance for components 1, 2, and 3 changes to 55.8 %, 12.7 %, and 10.5 %, respectively. This effect of the varimax rotation maximizes the loading of each variable on one of the extracted factors, while minimizing loadings on all other factors. As in the log-level analysis, we assumed that the factors should be independent. Therefore, we used an orthoginal rotation (i.e., varimax). Total extraction from the three PCs is 82.6 %. The reason for the change in variance explained by each factor is show in Table 26 where

Sweep and Length show a higher correlation with component 2 in the rotated solution than in the un-rotated solution. Additionally, the Live Crown Ratio shows a higher correlation with component 3 in the rotated solution than in the un-rotated solution. The two PC's – including size and form – found to be significant in the GLVR regression models are supported by other researchers, whom have found that individually measured variables, such as DBH, height and taper, had significant impact on GLVR, of which DBH was the strongest (Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a; Liu 2007b). In this study, DBH is strongly correlated to tree size, as is shown by the PCA in Section 4.3. Variable loadings to PC1 for DBH, height and taper are 0.932, 0.785 and 0.363, respectively. Thus, this supports previous work (Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a; Liu 2007b) that DBH, height and taper are significant in determining GLVR in this order.

Table 25. Total variance explained by components with extraction sum of squares loadings and rotated sum of squares loadings.

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.750	57.50	57.50	5.750	57.50	57.50	5.578	55.78	55.78
2	1.137	11.37	68.87	1.137	11.37	68.87	1.266	12.66	68.44
3	1.010	10.10	78.97	1.010	10.10	78.97	1.053	10.53	78.97
4	0.821	8.21	87.18						
5	0.663	6.63	93.81						
6	0.294	2.94	96.75						
7	0.129	1.29	98.04						
8	0.097	0.97	99.01						
9	0.059	0.59	99.60						
10	0.040	0.40	100.00						

The scree plot in Figure 22 also represents the contribution of each component in explaining the variance in the data. Figure 22 shows component 1 explains the majority of the variance in the data and the subsequent components tend to plateau. Generally, where the component plateau is the intuitive cut-off point (Field 2005), however, all factors with eigen values above 1 should be retained for PCA when using Kaiser's criterion (Kaiser 1974). Therefore, components 1, 2 and 3 were retained for further analysis.

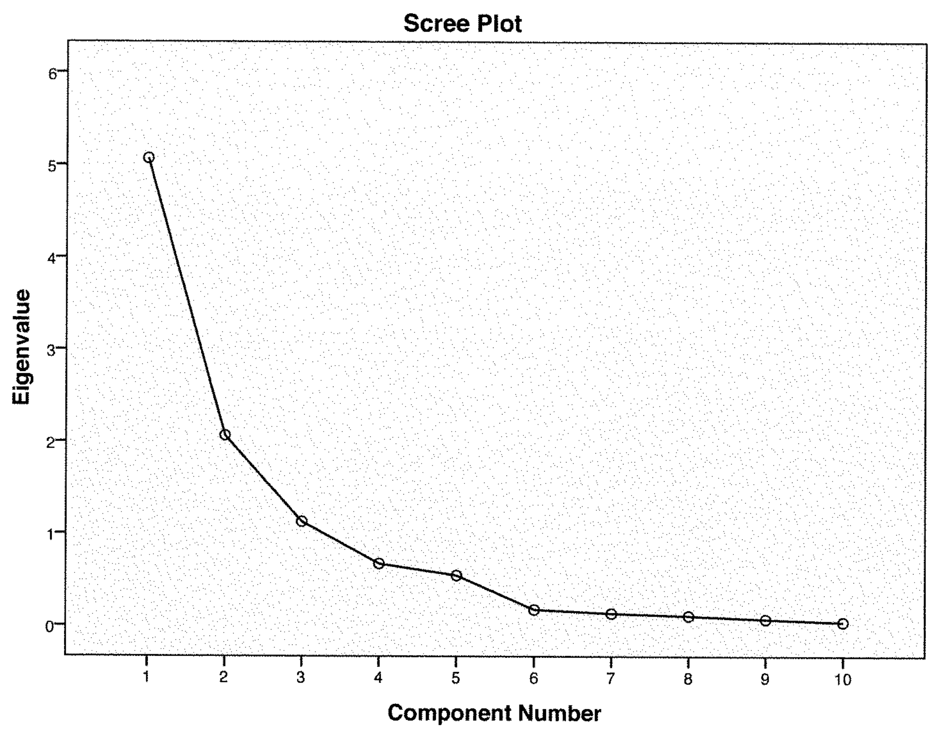


Figure 22. Scree plot showing eigenvalue trend by components.

The proportion of variance extracted from each tree variable is shown in Table 26. The average proportion of variance extracted is strong (0.790), with the highest value being 0.924 and the lowest value being 0.444 for d_10 (diameter at 10 m) and Taper, respectively. An average extraction value of 0.790 is well above the 0.6 level of acceptability (Kaiser 1974; Field 2005).Therefore, the PCA has a good level of extraction.

Table 26. Communalities in the variables explaining the proportion of data variance within each variable explained by the principal component factors.

	Communalities	
	Initial	Extraction
d_0.3	1	0.831
d_1.3	1	0.880
d_5	1	0.901
d_10	1	0.924
d_15	1	0.853
d_20	1	0.601
Ht_merch	1	0.693
LiveCrownRatio	1	0.919
Sweep_sum	1	0.850
Taper	1	0.444
Average	1	0.790

Extraction Method: Principal Component Analysis.

4.3.3 Factor Rotation and Interpretation

The Rotated Component Matrix (b) in Table 27 presents the clearer grouping of variables by component. Table 27 clearly shows that diameter and merchantable height measurements are strongly loaded on factor 1, which was also found by other researchers (Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a). Sweep, a type of form, is strongly loaded on factor 2. Live Crown Ratio, a descripter of tree branchiness (Kantola *et al.* 2007; Ikonen *et al.* 2009) is strongly loaded on PC 3. Taper has significant loadings on both PCs 1 and 2 and, therefore, is expressed as a combination of PC 1 (tree size) and PC 2 (tree form). The reason why taper is explained by a combination of PC1 and PC2 is due to the fact that taper is the diameter profile of the stem (i.e., tree size) and seeing that taper changes along the stem, it can also be described as tree form. These groupings of variables are shown in an un-rotated loading plot in Figure 23. A second loading plot is shown in Figure 24, where the plot is projected using a varimax rotation. The varimax rotation shows a clearer grouping of variables by PCs.

Table 27. Variable loadings by each principal component for the component matrix (a) and rotated component matrix (b) with values <0.3 (or >-0.3 when negative) excluded.

	Component Matrix(a)			Rotated Component Matrix(b)		
Variable	1	2	3	1	2	3
d_10	0.950	-	-	0.950	-	-
d_5	0.944	-	-	0.945	-	-
d_1.3	0.927	-	-	0.932	-	-
d_0.3	0.903	-	-	0.905	-	-
d_15	0.923	-	-	0.899	-	-
ht_merch	0.822	-	-	0.785	-	-
d_20	0.711	-	-	0.733	-	-
Sweep_sum	-	0.830	0.379	-	0.913	-
Taper	0.463	0.302	0.373	0.363	0.534	-
Live Crown Ratio	-	-0.537	0.793	-	-	0.958

(a) Extraction Method: Principal Component Analysis.
(b) Rotation Method: Varimax with Kaiser Normalization.

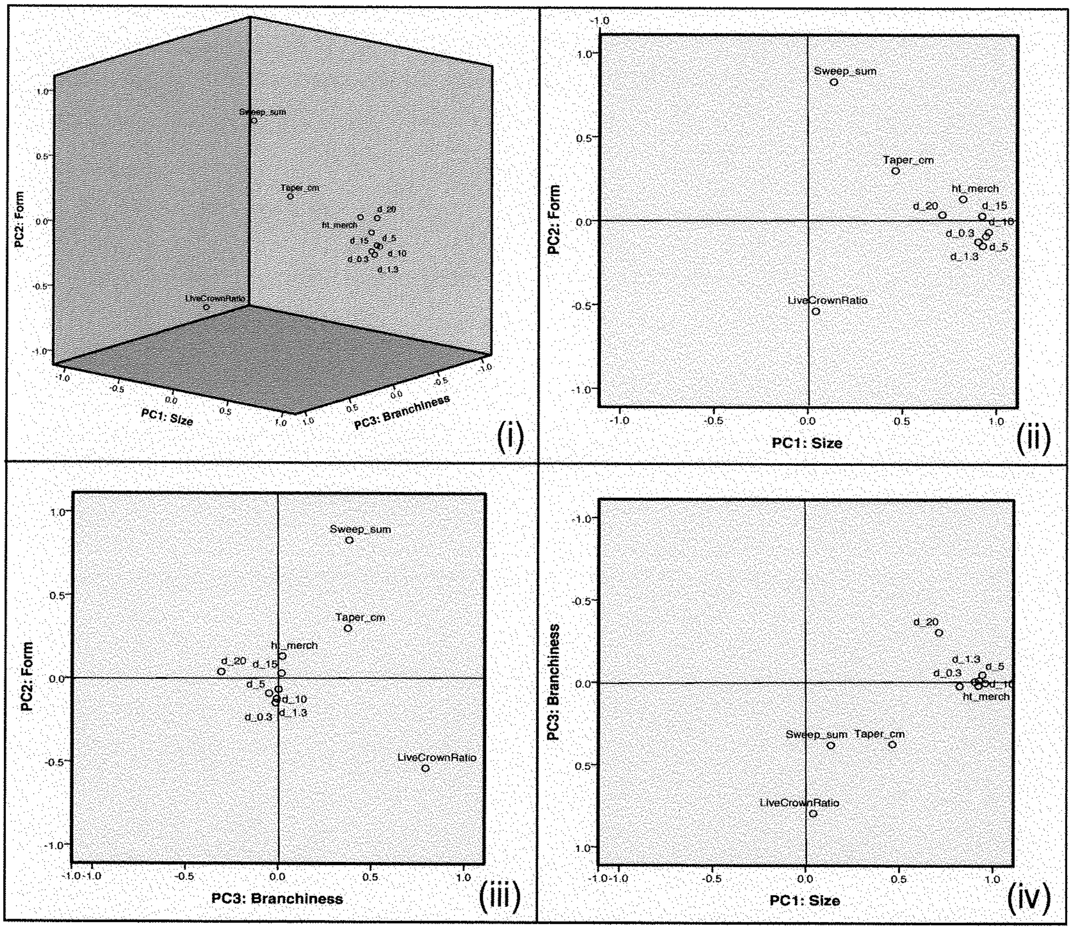


Figure 23. 3D and 2D plots of tree variables scores loaded by principal components (un-rotated solution).

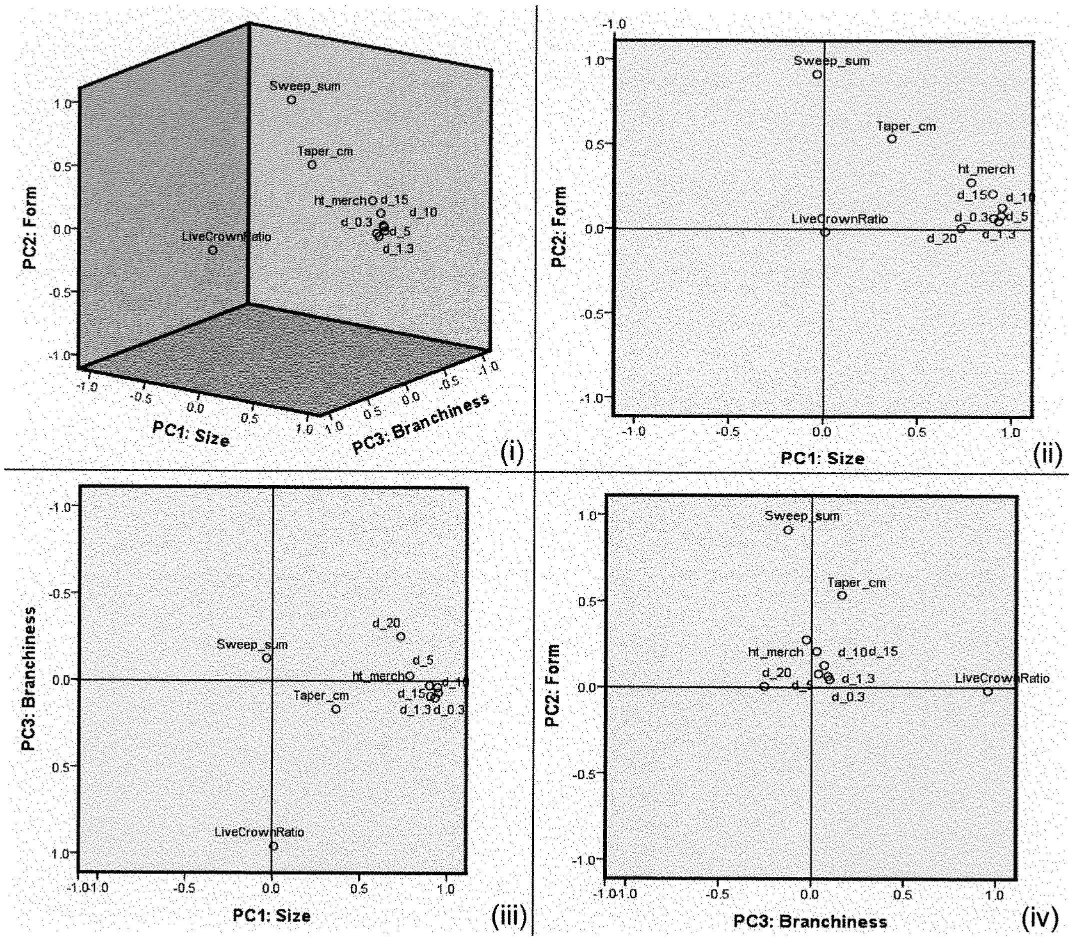


Figure 24. 3D and 2D plots of tree variables scores loaded by principal components (varimax rotated solution).

We can see from Figures 23 and 24 that once rotated, the variables loaded on axis 1, axis 2 and axis 3 clearly distinguish three names for the components. Component 1 can be termed Tree Size, Component 2 Tree Form and Component 3 Tree Branchiness. We see that PC1explains size components (tree merchantable length and diameters along stem). PC2 explains the form variable (sweep). PC3 explains the branchiness variable (Live Crown Ratio). The taper variable is explained by an interaction of size and form components.

Stem sweep– described by PC2 – has not been included in many lumber recovery valuation models due to its difficulty to model and requiring detail stem profile information (Eng *et al.* 1986; Gobakken 2000; Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a; Liu 2007b; Zakrzewski *et al.* 2010). However, other studies have found that stem sweep plays a critical role in determining a stem's GLVR (Kellogg and Warren 1984; Steele 1984; Middleton *et al.* 1989; Shi *et al.* 1990; Wagner and Taylor 1993; and Roos *et al.* 2000; Wilhelmsson and Moberg 2004; Moberg and Nordmark 2006). In the models that did not include stem sweep, the researchers generally did not have access to the sweep parameter in the dataset and mentioned that it would have played an important role in predicting GLVR (Eng *et al.* 1986; Gobakken 2000; Zakrzewski *et al.* 2010).

In addition to stem sweep, several studies have also found that branch and knot characteristics have a critical role in determining GLVR (Steele 1984; Middleton *et al.* 1989; Bharati and MacGregor 2003; Jones and Emms 2010). The reason why the analysis in this thesis did not find PC3 (branchiness) as significant in determining GLVR in a stem is due to knot characteristics not being recorded in the sawline log-scanning parameters at the sawmill. The log-level GLVR model presented in section 4.2, therefore, did not have branchiness or knot characteristics included in the log-level modeling used in the simulated bucking procedure outlined in section 3.4.1. This is certainly a short-coming in the log-and tree-level models produced in this thesis and the models should be applied with caution, as we know that knot (or branch) size, distribution and frequency along a log (or tree) is significant in determining final lumber value (Steele 1984; Middleton *et al.* 1989; Bharati and MacGregor 2003; Jones and Emms 2010).

4.4 TREE-LEVEL REGRESSION

4.4.1 Model Development

Table 28 presents the summary statistics for the 101 sample trees cruised in stands pre-harvest that were used in sawmill recovery study. Taking the results produced from the tree-level PCA analysis in Section 4.3, we conducted a regression analysis to determine if the principal components extracted from the tree level variables can be used to predict the GLVR from measured trees. Several models were created, tested and compared. Figure 25 shows the GLVR ($) by merchantable tree size (m^3) determined using the simulated bucking procedure and the log-level GLVR regression model described earlier (Model 7_L).

The use of PC's instead of the directly measured variables in the GLVR models was a different approach from most product recovery models done previously (Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a; Liu *et al.* 2007b). Often times, detailed stem profiles – as outlined in this study – are used as direct inputs to lumber recovery simulation software (Zhang and Tong 2005; Zhang *et al.* 2006: Liu *et al.* 2007a: Liu 2007b; Murphy *et al.* 2008). Due to cost and technical restraints, we were not able to capture the data in a way that was compatible with a product simulator. Instead, the multiple regression method was applied, and in order to eliminate the issue of collinearity between variables, the PCA method was then employed. The extracted PCs were used as independent variables in the regression models in order to allow for a higher level of data variance. Although this approach of using PC regression scores as independent variables has not been widely used to model GLVR, the conceptual theory to use a more detailed tree list of variables to describe tree variation in GLVR is supported by numerous researchers (Briggs 1989; Sachet *et al.* 1989; Beauregard *et al.* 2002; Wilhelmsson and Moberg 2004; Moberg and Nordmark 2006; Malinen *et al.* 2007; Murphy *et al.* 2008). The superior utility of using PC regression scores over 1 or 2 single measured variables is further supported by Bharati and MacGregor (2003) who found PCA regression scores superior to single-measured variables in the automated tracking of logs through the sawmill supply chain.

Table 28. Summary statistics of the data set used for establishing tree-level regression models (n=101).

	PC1: Size	PC2: Form	PC3: Branchiness	GLVR ($/tree)
Minimum	-2.39	-2.27	-2.44	11.16
Maximum	2.85	3.94	4.93	164.85
Mean	0.00	0.00	0.00	66.73
SD	1.00	1.00	1.00	28.65

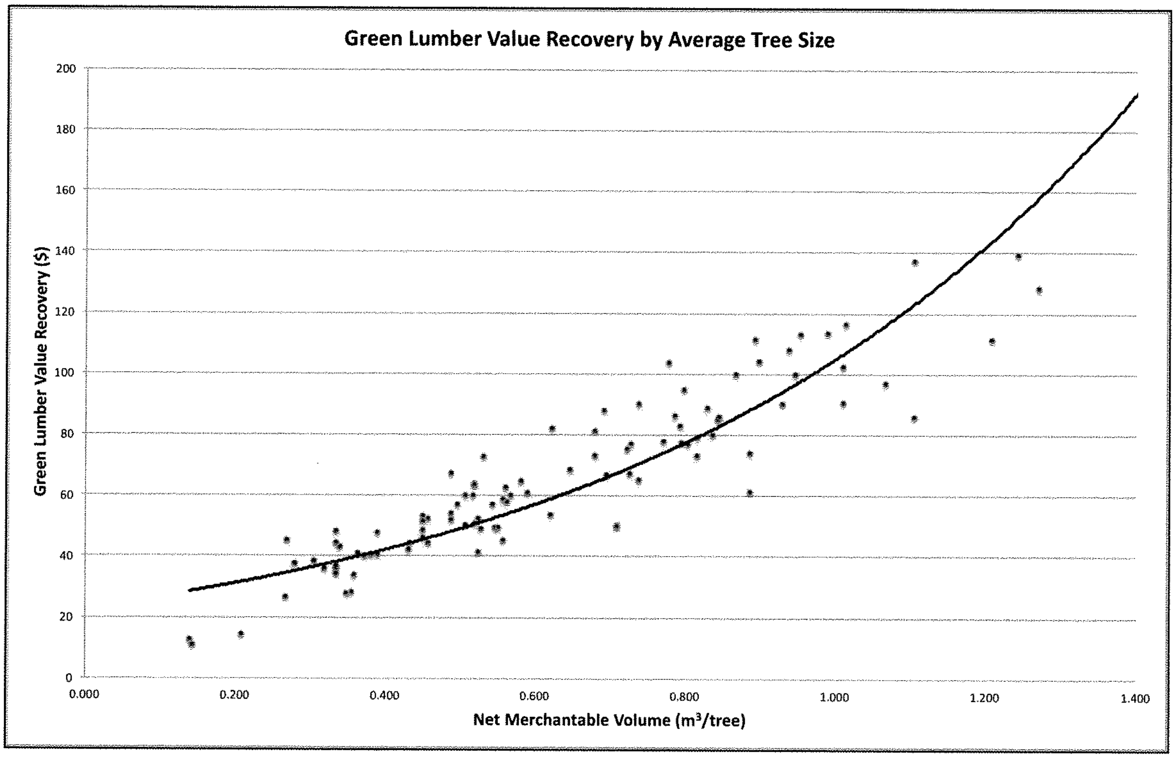


Figure 25. Model estimated GLVR ($/tree) by average tree size (net merchantable m^3/tree).

Table 29 shows the summarized results from the PCA analysis and how each PC is linearly correlated with GLVR. The linear correlation with the GLVR shows R^2 values of 0.938 with PC1, 0.001 with PC2 and <0.001 with PC3. From this preliminary look at correlations with the GLVR, it is apparent that GLVR has the strongest correlation with tree size and to a much lesser extent, tree form and tree branchiness. Figure 26 illustrates the strong fit of PC1 with GLVR. Figure 27 shows a poor linear fit between PC2 and the GLVR. Figure 28 shows no observable fit between PC3 and the GLVR, again this is due to knot characteristics not being measured. The ANOVA test of significance in Table 29, reveals that PC2:Form and PC3:Branchiness are not significant in prediction of GLVR. This is intuitive as the larger the tree size, the more lumber will be recovered. In other words with the same linear increase of tree size there is an exponential increase in recoverable value. With an increase in the tree form variability (an increase indicates poorer form), there is a corresponding drop in recoverable value. Branchiness has the lowest correlation because knot size was not taken into account at the sawline recovery phase upon which all log and tree-level models are built. Previous studies support the observation that tree size and form variables (diameter, height and taper) to be significant in GLVR of which tree size was most important (Zhang and Tong 2005; Zhang *et al.* 2006: Liu *et al.* 2007a).

Table 29. Summarized results of tree-level principal component analysis used in regression modeling.

		Principal Componenets Extracted		
		PC1: Size	PC2: Form	PC3: Branchiness
Eigenvalue		5.578	1.266	1.053
Proportion of total variance (%)		55.779	12.658	10.531
Loadings (>0.3)	d_10	0.950	-	-
	d_5	0.945	-	-
	d_1.3	0.932	-	-
	d_0.3	0.905	-	-
	d_15	0.899	-	-
	ht_merch	0.785	-	-
	d_20	0.733	-	-
	Sweep_sum	-	0.913	-
	Taper	0.363	0.534	-
	Live Crown Ratio	-	-	0.958
Linear R^2 predicting GLVR($)		0.938	0.001	<.001
Anova Sig. Test		<0.001	0.702	0.884

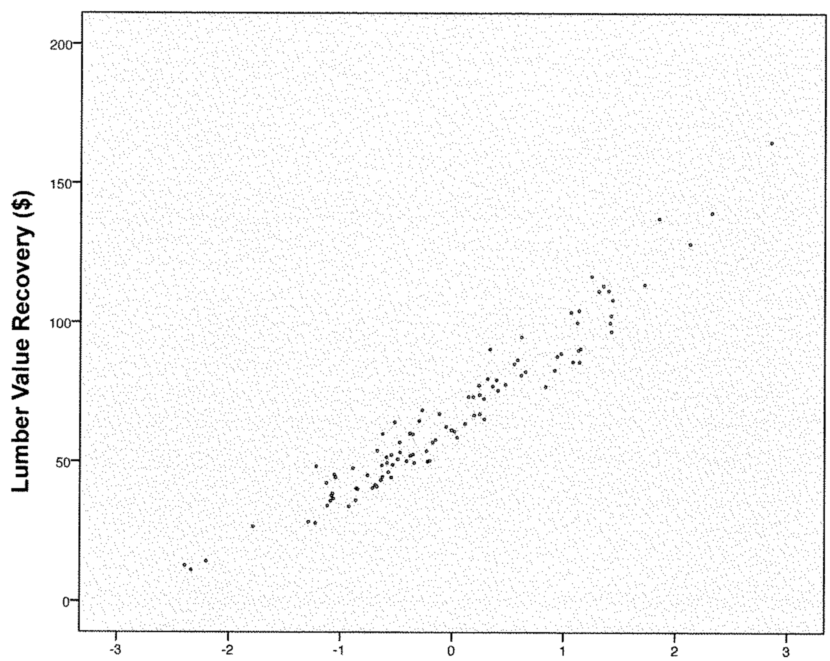


Figure 26. Plots of principal component 1 regression scores against GLVR ($/tree).

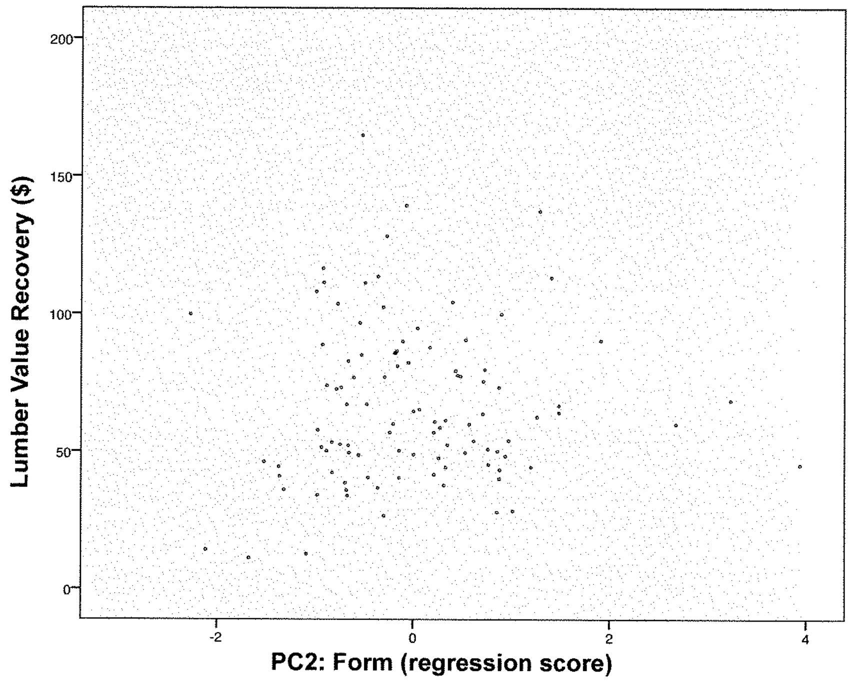


Figure 27. Plots of principal component 2 regression scores against GLVR ($/tree).

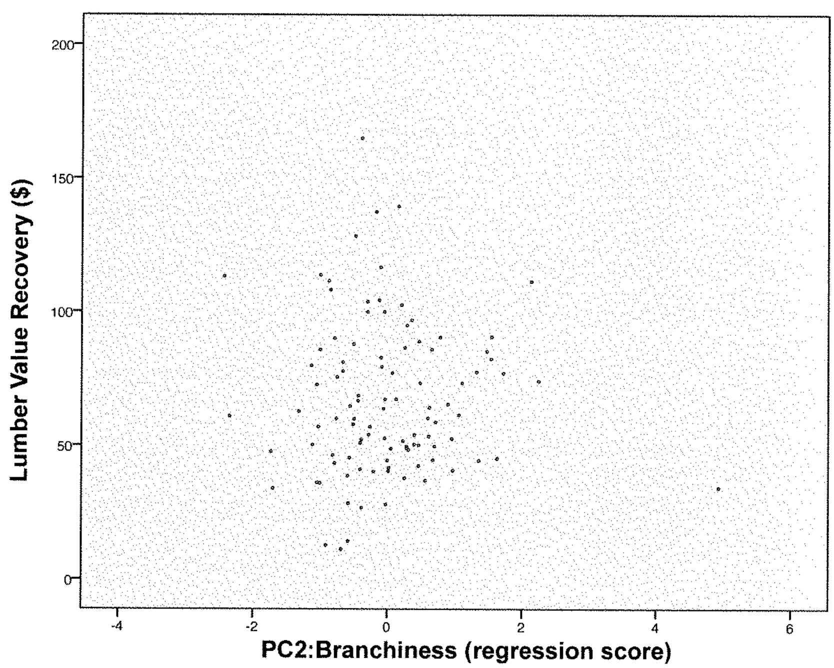


Figure 28. Plots of principal component 3 regression scores against GLVR ($/tree).

4.4.2 Model Comparison and Evaluation

Three types of models were tested in the analysis. These types were linear functions, non-linear power functions and non-linear exponential functions. Similar functions were tested on lumber recovery studies by other researchers (Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a; Liu *et al.* 2007b). A total of 8 individual models were run and the model forms are shown in Table 30 with the parameter estimates and criteria shown in Table 31. Residual plots are shown in Figures 28 and 29.

Table 30. Model forms for estimating lumber value recovery using tree-level principal components size (S), form (F), and branchiness (B): T is GLVR in $/tree, and a_0, a_1, a_2 and a_3 are constant coefficients.

Model number	Model form
1_T	$T = a_0 + a_1 * S$
2_T	$T = a_0 + a_1 * S + a_2 * F + a_3 * B$
3_T	$T = a_0 + a_1 * S + a_2 * F$
4_T	$T = a_0 + a_1 * S^2 + a_2 * F^2 + a_3 * B^2$
5_T	$T = a_0 + a_1 * S^3 + a_2 * F^3 + a_3 * B^3$
6_T	$T = a_0 + a_1 * \exp(a_2 * S)$
7_T	$T = a_0 + a_1 * \exp(a_2 * S + a_3 * F)$
8_T	$T = a_0 + a_1 * F + a_2 * \exp(a_3 * S)$

In deciding which models to use, an exploratory method was used as per other researchers (Zhang and Tong 2005; Zhang *et al.* 2006; Liu *et al.* 2007a). First a stepwise regression analysis was used to determine if combining PC1, PC2 and PC3 as predictor variables would improve the model. The results of model parameter estimates and criteria are shown in Table 31.

Three power functions were formulated and tested. Model 4_T employed a 2^{nd} order power function, which yielded the lowest R^2 value of 0.099 and the highest MSE value of 778.00. Model 5_T shows a third order power function with an improved fit compared to Model 4_T with an R^2 of 0.521 and a MSE 327.00. The observation that with higher order power functions performing better than lower order power function, supports the direction to use nonlinear models over linear (Field 2005). Neither power functions performed better than the linear functions, therefore exponential functions were explored.

Three exponential functions were compared. Liu *et al.* (2007a) also compared exponential functions in modeling lumber value recovery in black spruce using measured variables diameter and height. They found that exponential functions did not perform as well as polynomial functions. This difference in superior model forms between this thesis and Liu et al. (2007a) observations may be due to the use of different types of independent variables (i.e. PC regression scores rather than measured variables). Model 6_T included PC1 in an exponent function, yielding an R^2 of 0.951 and a MSE of 40.70. Model 7_T used PC1 and PC2 in the exponential function yielding an R^2 of 0.957 and a MSE of 36.74. Model 8_T kept PC1 as an exponential function and reduced PC2 to a linear function yielding an R^2 of 0.956 and a MSE of 37.17. Figure 30 illustrates the exponential model residual plots for the models created. Model 7_T performed only marginally better than Models 6_T and 8_T and there is no noticeable difference between residual plots. The reason for so little difference between models 6_T and 8_T is because of the extremely small affect PC2 (tree form) plays on the overall GLVR modeling that is mostly attributed to PC1 (tree size). Tree size was also found to be the most significant factor in lumber recovery modeling by other researchers (Kellogg and Warren 1984; Zhang *et al.* 2002; Zhang and Tong 2005; Liu *et al.* 2007a). All three exponential models performed better than the linear and power-function models thus an exponential function was selected to estimate the GLVR using tree-level data. An exponential function is a logical choice as the data plotted in Figure 25 shows an overall exponential data trend. It is important to note that the inclusion of PC2 only marginally improved the exponential models 7_T and 8_T showing that size is definitely the most important component in predicting the GLVR from the tree-level data. The observations that tree size, followed by tree form are the most important factors in determining GLVR is supported by previous studies (Kellogg and Warren 1984; Zhang *et al.* 2002; Zhang and Tong 2005; Liu *et al.* 2007).

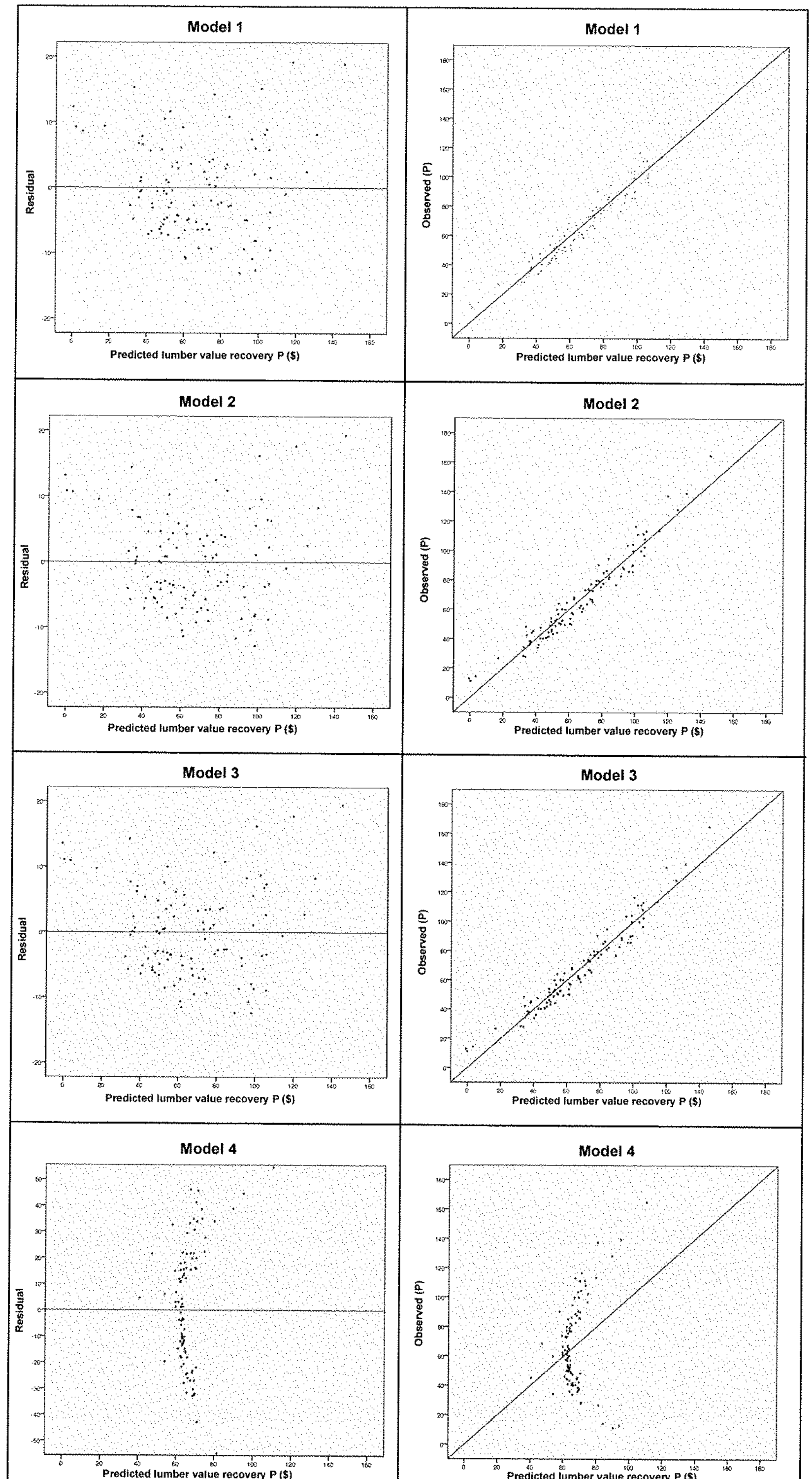


Figure 29. Plots of residuals against predicted green lumber value recovery P ($/tree) in the sawmill for tree-level data (Models 1 to 4).

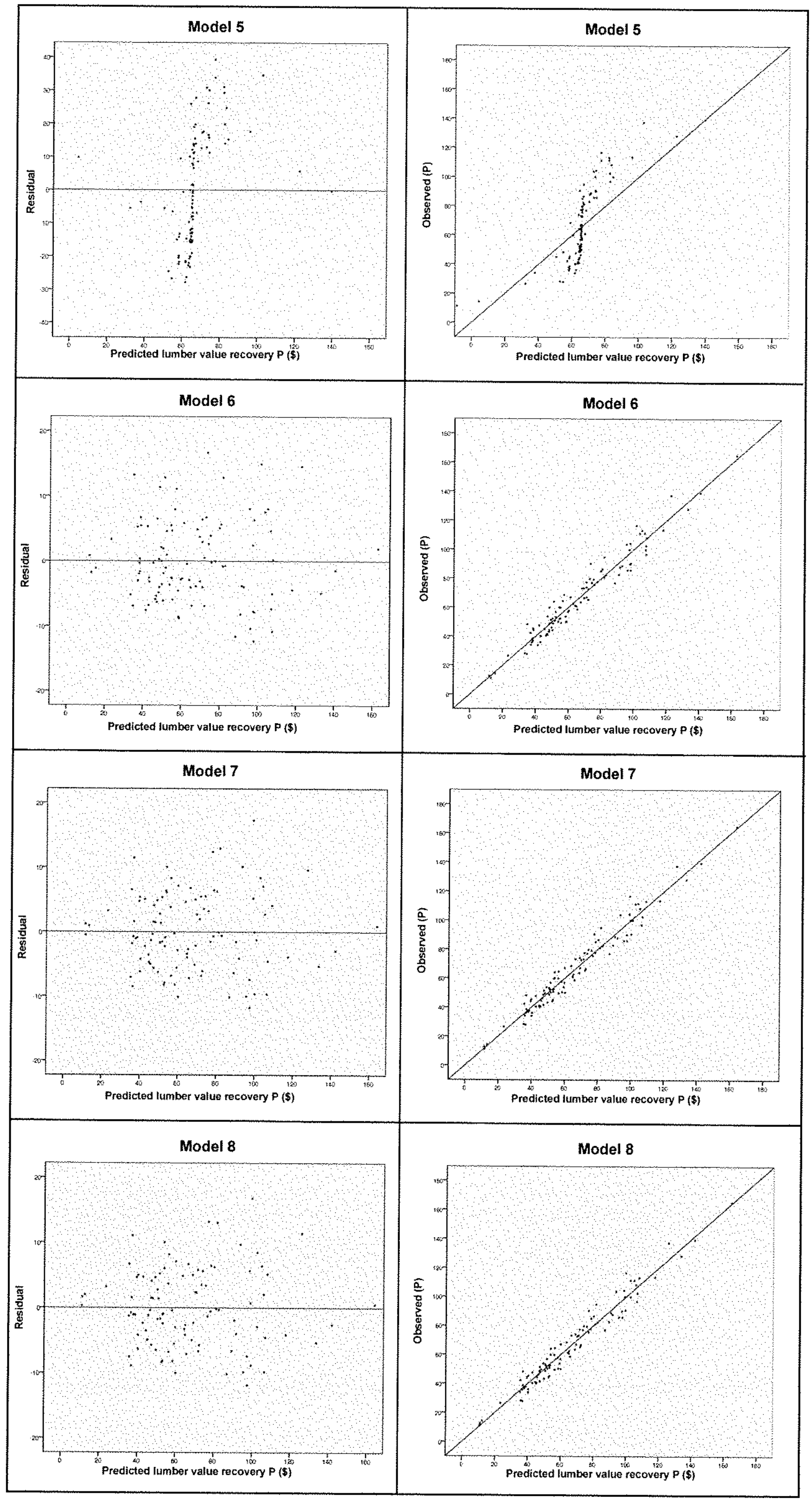


Figure 30. Plots of residuals against predicted green lumber value recovery P ($/tree) in the sawmill for tree-level data (Models 5 to 8).

Table 31. Parameter estimates and statistical criteria for the 8 tree-level regression models.

Model number	Parameters				Criteria		
	a_0	a_1	a_2	a_3	R^2	MSE	RMSE
1_T	66.73	0.969			0.938	51.239	5.811
2_T	66.73	0.969	0.039	-0.015	0.940	50.856	5.774
3_T	66.73	0.969	0.039		0.940	50.517	5.735
4_T	63.096	5.857	-1.568	-0.618	0.099	778.000	21.741
5_T	65.598	5.901	-0.192	-0.164	0.621	327.000	15.098
6_T	-91.606	156.014	0.171		0.951	40.702	5.092
7_T	-73.429	137.522	0.194	0.016	0.957	36.737	4.236
8_T	-72.814	2.006	136.923	0.193	0.956	37.171	4.925

It was found that the exponential function of PC1 and PC2 (Model 7_T) estimated GLVR the most accurately. However, inclusion of PC2 only slightly improved the exponential models, thus showing that accurate models can be generated with only PC1. PC3 negatively affected model prediction accuracy and was removed from further modeling trials.

$T_j = 73.429 + 137.522 * \exp(0.194*S_{PC1} + 0.016*F_{PC2})$ [Model 7_T]

Where,

T_j represents GLVR (CAN$) from a log j,

S_{PC1} denotes the regression score for PC1(Tree Size) extracted from the tree-level variables (diameter profile, merch. height, taper and sweep) in the PCA for each tree.

F_{PC2} denotes the regression score for PC2(Tree Form) extracted from the tree-level variables (diameter profile, merch. height, taper and sweep) in the PCA for each tree.

The lumber value estimation procedure for trees utilizes the PCA to extract PC regression scores from 11 external tree variables (diameter at 0.3 m, diameter at 1.3 m, diameter at 5 m, diameter at 10 m, diameter at 15 m, merch. height, taper and sweep). These regression scores of PC1:Size and PC2:Form are used as the dependent variables in the nonlinear exponential multiple regression formula (Model 7_T) to estimate tree-level GLVR. This formula was used to calculate block-level GLVR estimates.

Using the PCs over single measured variables (i.e., DBH and height) in the GLVR models have shown to better predict GLVR in conifer *spp*. For example, Zhang and Tong (2005), Zhang *et al.* (2006) and Liu *et al.* (2007a) have found DBH and total tree height capable of predicting up to 92% in lumber value creation. The study has shown that PC regression scores are capable of predicting 95.7% of the variation in GLVR; thus, demonstrating that the inclusion of a larger list of measured variables in initial inventory can yield a more accurate prediction. The increase in prediction accuracy is however marginal between this thesis model and models using DBH and height, and most likely is not worth the higher cost of data collection. In order for the methodology described in this study to become operationally feasible, more cost-effective data collection tools will need to be implemented; for example TLS, shown by Murphy *et al.* (2008 and 2010b).

4.5 BLOCK-LEVEL PRODUCTIVITY COMPARISON

Table 32 shows the block-level productivity summary produced for the three study sites. In the review of the literature none of the tree-level lumber recovery modeling methods provided block-level production summaries. The studies described in the literature were focused on generating tree level models and were not concerned with an applied case study of how these models can actually be applied to provide block-level production summaries. However, the utility of the tree-level recovery models to be used in operational planning were discussed (Kellogg and Warren 1984; Zhang *et al.* 2002; Zhang and Tong 2005; Liu *et al.* 2007a).

The production summary shows that volume of jack pine per hectare as well as its piece size, form and quality impact the overall green lumber value per cubic metre of jack pine per hectare ($/m^3 of net merchantable volume Pj/ha). Blocks 383, 559 and 568 have overall GLVR ($/m^3) of jack pine of $93.0, $78.6 and $74.7, respectively. For example, we see that the average piece size of block 568 is larger than block 559, yet the $/m^3 of jack pine in the block is higher for block 559. However, jack pine in block 568 had high defect severity, while block 559 had low defect severity. This comparison between block 559 and 568 demonstrates that stem form and quality, and not only stem size, are also important parameters to measure in determining a tree's (and block) profitability in sawmilling. The observations that tree size, form, and quality are accordingly significant in determining GLVR is supported by previous researchers who built and tested tree-level lumber recovery models (Kellogg and Warren 1984; Zhang *et al.* 2002; Zhang and Tong 2005; Liu *et al.* 2007a).

Table 32. Block-level production summary.

Block	383	559	568
All spp. net merch vol (m^3/ha)	235.6	202.7	256.8
Pj, net merch vol (m^3/ha)	169.9	133.3	96.2
Pj Tree-length Vol. (m^3/ha)	135.9	106.6	77.0
Pj, average net merch tree size (m^3)	0.7	0.4	0.5
Pj, average defect severity	Med	Low	High
Area harvested to deliver 500m^3 tree-length (ha)	3.7	4.7	6.5
Weigh scale volume (m^3)	541.4	538.2	523.6
Stick scale in log yard (m^3/ha)	579.8	n/a	n/a
Total boards produced (FBM)	150,256.2	126,728.5	122,691.6
Total estimate of the boards produced ($)	58,135.4	49,109.0	46,683.1
LRF (FBM trimmer/ Pj net merch m^3)	240.4	202.8	196.3
LRF (FBM trimmers/weighscale m^3)	277.5	235.5	234.3
LRF (FBMtrimmers /xy scanners m^3)	294.0	293.0	275.0
Pj, board value recovery ($/ha)	15,803.5	10,474.0	7,185.5
Pj, board recovery (FBM/ha)	40,845.6	27,028.7	18,884.7
Pj, board value recovery ($/m^3 of net merch Pj /ha)	93.0	78.6	74.7

5. CONCLUSIONS

5.1 SYNTHESIS

The purpose of this research project was to test whether an enhanced forest inventory can be integrated with sawmill production data to model GLVR prior to forest harvesting. A case study was conducted for a stud sawmill that tracked supply chain data from pre-harvest inventory through to lumber conversion. Over the 6 month period that data collection took place, we were able to compile a value chain dataset for a random-length sawmill in Thunder Bay measuring three sites from pre-harvest through to the GLVR. Additionally PCA and multiple regression analyses were conducted on log- and tree-level datasets to model the GLVR in the stud sawmill supply chain.

For the log-level analysis, we found that there were significant levels of similarity between the log-level variables measured. Second, out of the 7 log-level variables analyzed, 6 could be used to construct two principal components (Size and Form) that explain 80 % of the total log-level data variance. Third, we found that PC1:Size contributes the most in predicting the GLVR and PC2: Form, although significant, only made a minor improvement to the log-level regression model's performance. Finally, we found that the nonlinear exponential function provided the strongest prediction of the measured GLVR (Model 7_L).

$L_j = 27.467 + 0.0864*F_{PC2} + 36.626*\exp(0.172*S_{PC1})$ [Model 7_L]
Where,
L_j represents GLVR(CAN$) from a log j,
S_{PC1} denotes the regression score for PC1(Log Size) extracted from the log-level variables (small end diameter, large end diameter, volume, length, taper and sweep) in the PCA for each log.
F_{PC2} denotes the regression score for PC2(Log Form) extracted from the log-level variables (small end diameter, large end diameter, volume, length, taper and sweep) PCA for each log.

For the tree-level analysis, we found that there were significant levels of similarity between the tree-level variables measured. Second, out of the 29 tree-level variable analyzed, 10 could be used to extract three principal components (Size, Form and Branchiness) that explain 79 % of the total tree-level data variance. Third, we found that PC1:Size contributes the most in predicting the GLVR and PC2:Form, although statistically significant, only contributes a minor improvement to the models performance. PC3:Branchiness had a negative effect on model performance and was removed as a predictor variable. Finally, we found that the nonlinear exponential function created the strongest fit to the measured GLVR (Model 7_T). It is important to note that the scanners in the sawmill used to create the logs-to-lumber correlations did not scan for branchiness or knot size. This is probably the reason why PC3:Branchiness was not found to be significant at the tree-level regression analysis.

$T_j = 73.429 + 137.522 * \exp(0.194 * S_{PC1} + 0.016 * F_{PC2})$ [Model 7_T]

Where,
T_j represents GLVR(CAN$) from a log j,
S_{PC1} denotes the regression score for PC1(Tree Size) extracted from the tree-level variables (diameter profile, merch. height, taper and sweep) in the PCA for each tree.
F_{PC2} denotes the regression score for PC2(Tree Form) extracted from the tree-level variables (diameter profile, merch. height, taper and sweep) in the PCA for each tree.

5.2 RESEARCH SIGNFICANCE

The Canadian forest industry is moving through massive structural change and is improving its supply chain. The business model is increasingly focused on allocating wood to sawmills to meet changing size, form and quality specifications due to final product market demands. This new business model requires the industry to map the forest product supply chain from standing timber through to end-user products. The research presented in this thesis has been able to track supply chain data from pre-harvest inventory through the lumber conversion process for a stud sawmill in Thunder Bay, Ontario. Few studies have been conducted like this in Canada where full scale mill run recovery data can be linked directly back to the standing timber inventory and used to create and validate recovery prediction models. The research presented in this paper has presented a methodology integrating an enhanced forest inventory and log-scanning data using PCA and multiple regression analysis to model the GLVR prior to harvest. The research not only fills a knowledge gap in operations research for the Canadian forest industry, but also demonstrates a methodology that sawmills can use to create mill-specific lumber recovery models that can be used in production planning. The models created can be integrated into a geo-spatial modeling framework for use in tactical and operational planning. The ability to do tactical and operational planning hinges on the pre-harvest inventory to be similar to the variables collected in this study.

The possible extension of building a geodatabase tracking system – on multiple scales – could be an adapted systems design built into forest management software (e.g., GEREMA by Central Computer Services, or FPInterface by FPInnovations). This type of geodatabase system that maps specific product lines to unique forest attributes will allow managers to model their harvest allocations to fulfill specific orders while minimizing waste. Bringing GLVR modeling into the wood allocation process adds knowledge of volume and value into the decision-making. This knowledge will add a more precise forecasting ability to achieve specific value creation targets rather than simply volume. These types of forecasting tools support the paradigm shift in the forest industry of looking at the standing forest more as a warehouse with a precisely measured inventory.

5.4 FUTURE STUDIES

Future research opportunities include developing the GLVR models that use even fewer tree-level variables but tailored to specific sawmills (Zhang and Tong 2005). Using fewer tree-level variables in the modeling process has the potential to save inventory costs. However recent research is trending towards collecting more tree-level variables, not less, in modeling product recovery (Steele 1984; Middleton *et al.* 1989; Liu *et al.* 1989; Shi *et al.* 1990; Wagner and Taylor 1993; Roos *et al.* 2000; Zhang and Tong 2005; Liu *et al.* 2007a; Liu *et al.* 2007b). Collecting more tree-level variables is becoming more feasible with the advancement of Lidar technologies (Groot and Pitt 2010). Another future research opportunity would be to compare generic lumber recovery models (e.g. Zhang *et al.* 2001; Zhang and Tong 2005; Zhang *et al.* 2006; Liu 2007a; Liu 2007b) and the mill-specific models presented in this thesis to determine if indeed mill-specific models produce significantly different GLVR estimates from pre-harvest data. Another future research project could collect similar data for additional sites to validate the log- and tree-level GLVR models presented in this thesis.

The dataset collected in this research project could be used to predict lumber grade yields in jack pine using pre-harvest inventory variables using a similar modeling approach as Liu *et al.* (2007b). Locally validated models could be generated for specific sawmills to aid in forecasting lumber grade yields using a similar approach to this thesis. High-resolution imagery analysis (e.g. ADS-40) could also be used to generate stem profile data using regionally calibrated height-dbh curves and taper equations, and compare model performance with the cruising methodology explained in this thesis. Further, stem models and recovery models could be used, using terrestrial Lidar scanning equipment (similar to Murphy 2008) that has not yet been studied in Ontario's boreal forest. In order to have a holistic picture for the supply chain of the stud sawmill studied, there is need for a broader range of sites and stand types and ages, and expand to other species (i.e., black spruce, white spruce and balsam fir). Additionally, tracking the lumber through the planer mill and final grading process would greatly enhance the estimation of the final product value. Finally, additional studies could look at modeling product recovery using pre-harvest inventory to include more products such as chips, veneer quality wood, biomass, etc. and model a forest stand with multiple products transported to multiple forest products manufacturers.

LITERATURE CITED

Adams, T., and R.Y. Cavana. 2009. Systems Thinking in the Forestry Value Chain–A Case Study of the New Zealand Emissions Trading Scheme: Proceedings of the 53rd Annual Meeting of the International Society for the Systems Sciences, 12-17 July 2009, University of Queensland, Brisbane, Austrailia. 156pp.

Anon. 2007. Industry at a Crossroads: Choosing the Path to Renewal. Forest Products Association of Canada.

Anon. 2009. NSERC Forest Sector R&D Initiative. http://www.nserc-crsng.gc.ca/Professors-Professeurs/RPP-PP/SNGSuppForestry-SRSSuppforestier_eng.asp. December 6, 2010.

Anon. 2011. Natural Science and Engineering Research Institute of Canada. NSERC Forest Sector R&D. http://forest-foret.nserc-crsng.gc.ca/index_eng.asp. September 15, 2011.

Anon. 2011a. FPInnovations. FPInnovations Divisions. http://www.fpinnovations.ca/strength_e.htm. September 15, 2011.

Anon. 2011b. Forac. FORAC Consortium. http://www.fpinnovations.ca/strength_e.htm. September 17, 2011.

Anon. 2011c. Cirrelt. Mission. https://www.cirrelt.ca/Default.aspx?Page=MISSION. September 28, 2011.

Anon. 2012. Fpac. Forest Products Assocaition. http://www.fpac.ca/index.php/en/.ca. February 6, 2012

Aubry, Carol A, W T Adams, and Thomas D Fahey. 1998. Determination of Relative Economic Weights for Multitrait Selection in Coastal Douglas-fir. Canadian Journal of Forest Research 28 (8) (August): 1164-1170.

Avery, TE, and HE Burkhart. 1983. Forest Measurement. 3rd Edition. McGraW-Hill Company, New York 387.

Ayer Sachet, JK, DG Briggs, and RD Fight. 1989. Tree Value System: Users Guide. Notes: 52.

Ball, P., P. Albores, and J. Macbryde. 2004. Requirements for Modelling e-Business Processes. Production Planning & Control 15 (8): 776–785.

Barbour, R. J, S. Johnston, J. P Hayes, and G. F Tucker. 1997. Simulated Stand Characteristics and Wood Product Yields from Douglas-fir Plantations Managed for Ecosystem Objectives. Forest Ecology and Management 91 (2-3): 205–219.

Barbour, R. J, and R. M Kellogg. 1990. Forest Management and End-product Quality: a Canadian Perspective. Canadian Journal of Forest Research 20 (4): 405–414.

Beaudoin, D., L. LeBel, and J. Frayret. 2007. Tactical Supply Chain Planning in the Forest Products Industry Through Optimization and Scenario-based Analysis. Canadian Journal of Forest Research 37 (1) (January): 128.

Beauregard, Robert, Rado Gazo, and Roderick Ball. 2002. Grade Recovery, Value, and Return-To-Log for the Production of NZ Visual Grades (Cuttings and Framing) and Australian Machine Stress Grades. Wood and Fiber Science 34 (3) (July 1): 485-502.

Bharati, M. H., J. F. MacGregor, and W. Tropper. 2003. Softwood Lumber Grading Through On-line Multivariate Image Analysis Techniques. Industrial & Engineering Chemistry Research 42 (21): 5345–5353.

Briggs, David George, and Or.) Pacific Northwest Research Station (Portland. 1989. Tree Value System : Description and Assumptions. Portland, Or.: U.S. Dept. of Agriculture, Forest Service, Pacific Northwest Research Station. 24.

Buongiorno, J. 2003. The Global Forest Products Model: Structure, Estimation, and Applications. Academic Press. Pp. 300.

Cambiaghi, A. R., S. D'Amours, and M. Ronnqvist. 2008. Core Supply Chain Management Business Processes – A Literature-Based Framework Proposition. In 3rd World Conference on Production and Operations Management, 5-8 August 2008, Gakushin University, Tokyo, Japan. 970-985.

Carson, W.W., H.E. Andersen, S.E. Reutebuch, and R.J. McGaughey. 2004. LiDAR Applications in Forestry: An Overview. In Annual Aspres Conference Proceedings, 23.

Chiorescu, S., and A. Gronlund. 2004. The Fingerprint Approach: Using Data Generated by a 3D Log Scanner on Debarked Logs to Accomplish Traceability in the Sawmill's Log Yard. Forest Products Journal and Index 54 (12): 269–276.

Christopher, M. 1998. Relationships and Alliances: Embracing the Era of Network Competition. Gower Press, Hampshire, England. 272.

Cid, Y., J. Frayret, F. Léger, and A. Rousseau. 2008. Agent-Based Simulation and Analysis of Demand-Driven Production Strategies in the Lumber Industry. Int J Prod Res. 47(22): 6295-6319.

Cooper, M.C., D.M. Lambert, and J.D. Pagh. 1997. Supply Chain Management: More Than a New Name for Logistics. International Journal of Logistics Management, The 8 (1): 1–14.

D'Amours, S., R. Epstein, A. Weintraub, and M. Rönnqvist. 2011. Operations Research in Forestry and Forest Products Industry. Wiley Encyclopedia of Operations Research and Management Science: 27.

D'Amours, S., E. Gunn, and R. Pulkki. 2010. NSERC VCO Network. Optimizing the Modern Forest Bioeconomy Networks. In ECANUSA Forest Science Conference, 16. Edmundston, New Brunswick.

D'Amours, S., M. Rönnqvist, and A. Weintraub. 2008. Using Operational Research for Supply Chain Planning in the Forest Products Industry. Infor 46 (4) (November): 265.

Davis, L., and K. Johnson. 1987. Forest Management. Mcgraw-Hill College. 816.

Deadman, MW. 1990. MicroMARVL Pre-harvest inventory–User Guide. NZ Forest Research Institute, Software Series 7: 13.

Deadman, MW, and CJ Goulding. 1979. A Method for Assessment of Recoverable Volume by Log Types. New Zealand Journal of Forestry Science 9 (2): 225-239.

Dowding, B., and G. Murphy. 2010. Estimating Spatial Changes in Acoustic Velocity in Felled Douglas-fir Stems . Theses, Dissertations and Student Research Papers (Sustainable Forest Management, Forest Engineering, & Forest Management).

Dramm, J.R. 2004. Log Sort Yard Economics, Planning, and Feasibility. US Dept. of Agriculture, Forest Service, Forest Products Laboratory 146: 35.

Ellram, L.M. 1991. Supply-Chain Management: The Industrial Organisation Perspective. International Journal of Physical Distribution & Logistics Management 21 (1): 13–22.

Emmett, B. 2005. Focus on: Fibre for a Forest Fix. ADM Canadian Forest Service in CFS Viewpoint. 12.

Eng, G., HG Daellenbach, and AGD Whyte. 1986. Bucking Tree-length Stems Optimally. Canadian Journal of Forest Research 16 (5): 1030–1035.

Eriksson, L. 1999. Introduction to Multi-and Megavariate Data Analysis Using Projection Methods (PCA & PLS). Umetrics AB. 289.

Feng, Y., S. D'Amours, and R. Beauregard. 2008. The Value of Sales and Operations Planning in Oriented Strand Board Industry with Make-to-order Manufacturing System: Cross Functional Integration Under Deterministic Demand and Spot Market Recourse. International Journal of Production Economics 115 (1): 189-209.

Fleischmann, B., and H. Meyr. 2004. Customer Orientation in Advanced Planning Systems. Supply Chain Management and Reverse Logistics. Springer, Berlin: 297–321.

Flodin, J., J. Oja, and A. Gronlund. 2008. Fingerprint Traceability of Logs Using the Outer Shape and the Tracheid Effect. Forest Products Journal 58 (4): 21.

Forget, P., S. D'Amours, and J. Frayret. 2008a. Multi-behavior Agent Model for Planning in Supply Chains: An Application to the Lumber Industry. Robotics and Computer Integrated Manufacturing 24 (5): 664-679.

Forget, P., T. Monteiro, S. D'Amours, and J. Frayret. 2008b. Collaborative Agent-based Negotiation in Supply Chain Planning Using Multi-behaviour Agents. CIRRELT, 2008-54. 21

Frayret, J., S. D'Amours, B. Montreuil, and L. Cloutier. 2001. A Network Approach to Operate Agile Manufacturing Systems. International Journal of Production Economics 74 (1-3): 239-259.

Frayret, J., S. D'amours, A. Rousseau, S. Harvey, and J. Gaudreault. 2007. Agent-based Supply-chain Planning in the Forest Products Industry. International Journal of Flexible Manufacturing Systems 19 (4): 358.

Ganeshan, R., and T.P. Harrison. 1995. An Introduction to Supply Chain Management. Penn State University, The United States. accessed at http://lcm.csa.iisc.ernet.in/scm/supply_chain_intro.html. July 13 2010.

Gaudreault, J., P. Forget, J. Frayret, A. Rousseau, and S. D'Amours. 2009. Distributed Operations Planning in the Lumber Supply Chain: Models and Coordination. International Journal of Industrial Engineering 17 (3): 19.

Gertler, Meric S., David A. Wolfe, and David Garkut. 2000. No Place Like Home? The Embeddedness of Innovation in a Regional Economy. Review of International Political Economy 7 (4): 688-718.

Gobakken, T. 2000. The Effect of Two Different Price Systems on the Value and Cross-cutting Patterns of Norway Spruce Logs. Scandinavian Journal of Forest Research 15 (3): 368–377.

Gordon, A., and D. Baker. 2004. Using External Stem Characteristics for Assessing Log Grade Yield: Comparing Stem Description and Stem Coding. In AusTimber 2004 Conference, Albury, NSW, Australia, 30–31.

Gordon, AD, and ME Lawrence. 1995. External Stem Quality assessment-MARVL. D. Hammond Ed: 190–191.

Goulding, CJ. 2000. The Forest as a Warehouse. United States Department of Agriculture Forest Service General Technical Report Nc: 276–282.

Goulding, CJ, CM Trotter, BK Hock, and S. Hitchcock. 2000. Determining the Location of Trees and Their Log Products Within a Stand. New Zealand Journal of Forestry 45: 34–39.

Grondin, F., and N. Drouin. 1998. Optitek Sawmill Simulator-User's Guide. Forintek Canada Corporation, Québec, Canada. 38.

Groot, A., and D. Pitt. 2010. Incorporating Fibre Attributes into Canadian Forest Inventories. In Proceedings of the Precision Forestry Symposium, 1-3 March 2010, Stellenbosch, South Africa. 18.

Guddanti, S., and S. J Chang. 1998. Replicating Sawmill Sawing with TOPSAW Using CT Images of a Full-length Hardwood Log. Forest Products Journal 48 (1): 72–75.

Gujarati, D., M, Damodar. 2002. Basic Econometrics. 4th ed. McGraw-Hill/Irwin. 1002.

Haartveit, E.Y., R.A. Kozak, and T.C. Maness. 2004. Supply Chain Management Mapping for the Forest Products Industry: Three Cases from Western Canada. Journal of Forest Products Business Research Volume 1 (5): 1.

Hillier, F.S., G.J. Lieberman, and M. Hillier. 1990. Introduction to Operations Research. Vol. 6. McGraw-Hill New York, NY. 1088.

Ikonen, V. P, S. Kellomäki, and H. Peltola. 2003. Linking Tree Stem Properties of Scots Pine (Pinus Sylvestris L.) to Sawn Timber Properties Through Simulated Sawing. Forest Ecology and Management 174 (1-3): 251–263.

Jensen, Anders. 2009. Valuation of Non-timber Forest Products Value Chains. Forest Policy and Economics 11 (1) (January 1): 34-41.

Jones, T. G, and G. W Emms. 2010. Influence of Acoustic Velocity, Density, and Knots on the Stiffness Grade Outturn of Radiata Pine Logs. Wood and Fiber Science 42 (1): 1–9.

Kantola, A., S. Harkonen, H. Makinen, and A. Makela. 2008. Predicting Timber Properties from Tree Measurements at Felling: Evaluation of the RetroSTEM Model and TreeViz Software for Norway Spruce. Forest Ecology and Management 255 (8-9): 3524–3533.

Kellogg, R. M, and W. G Warren. 1984. Evaluating Western Hemlock Stem Characteristics in Terms of Lumber Value. Wood and Fiber Science 16 (4): 583–597.

Lazar, A. 2007. Making Canada World-competitive: The Forestry Industry as a Case Study. Policy Options-montreal- 28 (7): 50.

Li, Y. 2009. Towards Small-footprint Airborne LiDAR-assisted Large Scale Operational Forest Inventory. Ph.D. Dissertation. University of Washington. 118.

Liu, C. M, W. A Leuschner, and H. E Burkhart. 1989. A Production Function Analysis of Loblolly Pine Yield Equations. Forest Science 35 (3): 775–788.

Liu, C., and S. Y. Zhang. 2005. Models for Predicting Product Recovery Using Selected Tree Characteristics of Black Spruce. Canadian Journal of Forest Research 35 (4): 930–937.

Liu, C., S. Y. Zhang, A. Cloutier, and T. Rycabel. 2007a. Modeling Lumber Value Recovery in Relation to Selected Tree Characteristics in Black Spruce Using the Optitek Sawing Simulator. Forest Products Journal 57 (4): 57.

Liu, C., S. Y. Zhang, and Z. H. Jiang. 2007b. Models for Predicting Lumber Grade Yield Using Tree Characteristics in Black Spruce. Forest Products Journal 57 (1/2): 60.

MacKenzie, J., and G. Bruemmer. 2009. Enhancing Canada's Forest Fibre. The Forestry Chronicle 85 (3): 353–354.

Malinen, J., M. Maltamo, and E. Verkasalo. 2003. Predicting the Internal Quality and Value of Norway Spruce Trees by Using Two Non-parametric Nearest Neighbor Methods. Forest Products Journal 53 (4): 85–94.

Maltamo, M., J. Malinen, P. Packalén, A. Suvanto, and J. Kangas. 2006. Nonparametric Estimation of Stem Volume Using Airborne Laser Scanning, Aerial Photography, and Stand-register Data. Canadian Journal of Forest Research 36 (2) (February): 426.

Mandel-Campbell, A. 2007. Why Mexicans Don't Drink Molson : Rescuing Canadian Business from the Suds of Global Obscurity. Vancouver: Douglas & McIntyre. 328.

Mentzer, J.T., W. DeWitt, J.S. Keebler, S. Min, N.W. Nix, C.D. Smith, and Z.G. Zacharia. 2001. Defining Supply Chain Management. Journal of Business Logistics 22 (2): 1–25.

Meredith Smith, J. 1999. Item Selection for Global Purchasing. European Journal of Purchasing & Supply Management 5 (3-4): 117–127.

Middleton, G. R. (Gerry), and S. Y. (Tony) Zhang. 2009. Characterizing the Wood Attributes of Canadian Tree Species: A Thirty-year Chronicle. The Forestry Chronicle 85 (3): 392-400.

Middleton, G. R., S. Bowe, J. Oja, D. Verret, and B. D. Munro. 2003. Utilizing CT Log Scanning to Add Value to British Columbia's Forest Estate: Enabling Software - Phase I. 26654 East Mall Vancouver, BC: Forintek Canada Corp. Western Division. R2003-0135. 85.

Middleton, G. R., L. A. Jozsa, L. C. Palka, B. D. Munro, and P. Sen. 1995. Lodgepole Pine Product Yields Related to Differences in Stand Density. Forintek Canada Corp., Western Laboratory. S5(01): 4.

Moberg, Lennart, and Urban Nordmark. 2006. Predicting Lumber Volume and Grade Recovery for Scots Pine Stems Using Tree Models and Sawmill Conversion Simulation. Forest Products Journal 56 (4) (April): 68.

Montreuil, B., J. Frayret, and S. D'Amours. 2000. A Strategic Framework for Networked Manufacturing. Computers in Industry 42 (2-3): 299-317.

Murphy, G. 2008. Determining Stand Value and Log Product Yields Using Terrestrial Lidar and Optimal Bucking: a Case Study. Journal of Forestry 106 (6): 317–324.

Murphy, G., and M. Acuna. 2008. Determining Stand Value and Log Product Yields Using Terrestrial Lidar and Optimal Bucking: a Case Study. Journal of Forestry 106 (6): 317–324.

Murphy, G., J. Lyons, M. O'Shea, G. Mullooly, E. Keane, and G. Devlin. 2010a. Management Tools for Optimal Allocation of Wood Fibre to Conventional Log and Bio-energy Markets in Ireland: a Case Study. European Journal of Forest Research 129 (6): 1057–1067.

Murphy, G., M. Acuna, and C. Dumbrell. 2010b. Tree Value and Log Product Yield Determination in Radiata Pine (Pinus Radiata) Plantations in Australia: Comparisons of Terrestrial Laser Scanning with a Forest Inventory System and Manual Measurements. Canadian Journal of Forest Research 40 (11): 2223–2233.

Nieuwenhuis, M. 2002. The Development and Validation of Pre-harvest Inventory Methodologies for Timber Procurement in Ireland. Silva Fennica 36 (2): 535–547.

Pilkerton, S.J. 2009. Thinning Aged Douglas-fir: An Analysis of Mobilization Costs and a Log Bucking Strategy for Revenue Improvement. Oregon State University, Oregon, USA. 282.

Pnevmaticos, S. M., Y. Corneau, and R. C. Kerr. 1980. Yield and Productivity in Processing Tree-length Softwoods. Mills in Quebec. Technical Report, Eastern Forest Products Laboratory, Forintek Canada Corp. (507E): 11.

Poirier, C.C. 1999. Advanced Supply Chain Management: How to Build a Sustained Competitive Advantage. Berrett-Koehler Publishers, San Francisco, CA, USA. 154.

Pulkki, R. 2004. Role of Supply Chain Management in the Wise Use of Wood Resources. Southern Forests: a Journal of Forest Science 191: 89–96.

Pulkki, RE. 1991. A Literature Synthesis on the Effects of Wood Quality in the Manufacture of Pulp and Paper. Forest Engineering Research Institute of Canada. TN-171: 10.

Rönnqvist, M. 2003. Optimization in Forestry. Mathematical Programming 97 (1): 267–284.

Roos, A., M. Flinkman, A. Jappinen, and M. Warensjo. 2000. Adoption of Value-adding Processes in Swedish Sawmills. Silva Fennica 34 (4): 423–430.

Ruel, J.C., A. Achim, R.E. Herrera, and A. Cloutier. 2010. Relating Mechanical Strength at the Stem Level to Values Obtained from Defect-free Wood Samples. Trees-Structure and Function 24(6): 1127-1135.

Shi, R., P. H Steele, and F. G Wagner. 1990. Influence of Log Length and Taper on Estimation of Hardwood BOF Position. Wood and Fiber Science 22 (2): 142–148.

Sjostrom, K., and L. O. Rask. 2001. Supply Chain Management for Paper and Timber Industries: Proceedings of the 2nd World Symposium on Logistics in [the] Forest Sector, 12-15 August 2001, Vaxjo, Sweden. 260.

Slater, S., and J. Narver. 1998. Customer-Led and Market-Oriented: Let's Not Confuse the Two. Strategic Management Journal 19 (10) (October): 1001-1006.

Steele, P. H. 1984. Factors Determining Lumber Recovery in Sawmilling. United States Department of Agriculture. Forest Service TR FPL-39. 10.

Tomppo, E., C. Goulding, and M. Katila. 1999. Adapting Finnish Multi-source Forest Inventory Techniques to the New Zealand Preharvest Inventory. Scandinavian Journal of Forest Research 14 (2): 182–192.

Tong, QJ, and SY Zhang. 2005. Impact of Initial Spacing and Precommercial Thinning on Jack Pine Tree Growth and Stem Quality. The Forestry Chronicle 81 (3): 418–428.

Uusijärvi, R. 2000. Automatic Tracking of Wood. Doctoral thesis. Superseded Deapartments, Production Systems. Sweedish University of Agricultural Sciences. Sweden. 170.

Uusitalo, J. 1997. Pre-harvest Measurement of Pine Stands for Sawing Production Planning. Acta Forestalia Fennica (259): 56.

Uusitalo. 2005. A Framework for CTL Method-based Wood Procurement Logistics. International Journal of Forest Engineering 16 (2): 37-46.

Uusitalo, J., and J. Isotalo. 2005. Predicting Knottiness of Pinus Sylvestris for Use in Tree Bucking Procedures. Scandinavian Journal of Forest Research 20 (6): 521-533.

Velde, D., J. Rushton, K. Schreckenberg, E. Marshall, F. Edouard, A. Newton, and E. Arancibia. 2006. Entrepreneurship in Value Chains of Non-timber Forest Products. Forest Policy and Economics 8 (7): 725-741.

Via, B. K, T. F Shupe, L. H Groom, M. Stine, and C. L So. 2003. Multivariate Modelling of Density, Strength and Stiffness from Near Infrared Spectra for Mature, Juvenile and Pith Wood of Longleaf Pine (Pinus Palustris). Journal of Near Infrared Spectroscopy 11 (5): 365–378.

Wagner, F. G, J. A Brody, D. S Ladd, and J. S Beard. 1996. Sawtimber Valuation and Sawlog Allocation Through Simulation of Temple-Inland Sawmills. Interfaces: 26(6):3–8.

Wagner, F. G., and F. W. Taylor. 1993. Low Lumber Recovery at Southern Pine Sawmills May Be Due to Misshapen Sawlogs. Forest Products Journal 43 (3): 53–55.

Wilhelmsson, L, and L Moberg. 2004. Predictions of Green Density of Log Assortments by Prediction Models and Industry Scaling. Skogforsk Report 569:35.

Wolfe, D. A., and M. S. Gertler. 2004. Clusters from the Inside and Out: Local Dynamics and Global Linkages. Urban Studies 41 (5/6): 1071-1093.

Zakrzewski, W, F Schnekenburger, and P Kozlowski. 2010. Tools for Optimizing Timber Product Mix: User's Guide for Visualizer-Buck. Forest Research Information Paper, OFRI 174: 30.

Zhang, S. Y, and Q. J Tong. 2005. Modeling Lumber Recovery in Relation to Selected Tree Characteristics in Jack Pine Using Sawing Simulator Optitek. Annals of Forest Science 62 (3): 219–228.

Zhang, S. Y., Chuangmin Liu, and Z. H. Jiang. 2006. Modeling Product Recovery in Relation to Selected Tree Characteristics in Black Spruce Using an Optimized Random Sawing Simulator. Forest Products Journal 56 (11/12) (December): 93.

Zhang, T., and JF Gingras. 1999. Twig Tweaking: Timber Management for Wood Quality and End-product Value. Canadian Forest Industries Nov./Dec, 1999: 43–46.